THE DERIVATION OF THE LINEAR BOLTZMANN EQUATION FROM A RAYLEIGH GAS PARTICLE MODEL

KARSTEN MATTHIES, GEORGE STONE, AND FLORIAN THEIL

Abstract. A linear Boltzmann equation is derived in the Boltzmann-Grad scaling for the deterministic dynamics of many interacting particles with random initial data. We study a Rayleigh gas where a tagged particle is undergoing hard-sphere collisions with background particles, which do not interact among each other. In the Boltzmann-Grad scaling, we derive the validity of a linear Boltzmann equation for arbitrary long times under moderate assumptions on higher moments of the initial distributions of the tagged particle and the possibly non-equilibrium distribution of the background. The convergence of the empiric dynamics to the Boltzmann dynamics is shown using Kolmogorov equations for associated probability measures on collision histories.

1. Introduction

The derivation of continuum equations from atomistic particle models is currently a major problem in mathematical physics with origins in Hilbert's Sixth Problem. A particular interest in this area is the derivation of the Boltzmann equation from atomistic particle dynamics. The first major work in this area was by Lanford [17] which showed convergence from a hard-sphere particle model for short times by using the Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy, see e.g. [8, 11, 33]. A recent major work by Gallagher, Saint-Raymond and Texier [14] continued the BBGKY development and proved convergence to the Boltzmann equation for short times for both hard-sphere and short-range potentials. The latter of which was improved by Pulvirenti [29]. A particular difficulty has been proving convergence for arbitrarily long times, since the aforementioned results hold only for times up to a fraction of the mean free flight time. In [5] Bodineau, Gallagher and Saint-Raymond were able to utilise the tools from [14] to prove convergence from a hard-sphere particle model to the linear Boltzmann equation for arbitrary long times in the case that the initial distribution of the background is near equilibrium. They were further able to use the linear Boltzmann equation as an intermediary step to prove convergence to Brownian motion. Further in a following paper [6] the authors were able to consider weaker assumptions on the initial data and prove convergence to the Stokes-Fourier equations. For a general overview of the Boltzmann equation and the BBGKY hierarchy we refer to the books [10, 11, 32].

A new method to tackle this problem has recently been developed in a series of papers [24, 25, 26]. This method employs semigroup techniques to study the evolution of collision trees rather than the BBGKY hierarchy. This comes from studying the distribution of the history of the particles up to a certain time rather than the distribution of the particles at a specific time. These papers have been able to prove convergence for arbitrary times but only for a simplified particle interaction system. This is the approach that we continue to develop in this paper.

1.1. The BBGKY Hierarchy.

The standard approach to tackling the derivation of the Boltzmann equation is via the BBGKY hierarchy. We refer specifically to [5] for hard spheres. If $f_N(t)$ represents the N particle distribution resulting from hard sphere dynamics at time t then, away from collisions, f_N satisfies the Liouville equation,

$$\partial_t f_N(t) + v \cdot \partial_x f_N(t) = 0.$$

Considering this equation in weak form and integrating we have, away from collisions,

$$\partial_t f_N^{(s)}(t) + \sum_{1\le i\le s} v_i \cdot \nabla_{x_i} f_N(t) = C_{s,s+1} f_N^{(s+1)}(t), \tag{1.1}$$

for $s = 1, \dots, N$ where $f_N^{(s)}$ denotes the s particle marginal and where $C_{s,s+1}$ is the effect on the distribution of the first s particles by a collision with another particle given explicitly in (2.6) below. The system of N equations (1.1) is known as the BBGKY hierarchy. If the initial distribution of the N particles is given by,

$$f_N(0) = \frac{1}{\mathcal{Z}_N} \mathbb{1}_{\mathrm{no}} f_0^{\otimes N},$$

where $\mathbb{1}_{\mathrm{no}}$ conditions on no initial overlap and $\mathcal{Z}_N$ is a normalising constant, then the initial distribution of $f_N^{(s)}$ is given by,

$$f_N^s(0, Z_s) = \int f_N(0, Z_n)\, \mathrm{d}z_{s+1} \dots \mathrm{d}z_N.$$

After successive time integration of (1.1) one obtains a representation in the form of a finite sum

$$f_N^{(s)}(t) = \sum_{k=0}^{N-s} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} \mathbf{T}_s(t-t_1)\mathcal{C}_{s,s+1}\mathbf{T}_{s+1}(t_1-t_2)\mathcal{C}_{s+1,s+2}\dots \\ \dots \mathbf{T}_{s+k}(t_k) f_N^{(s+k)}(0)\, \mathrm{d}t_k \dots \mathrm{d}t_1, \tag{1.2}$$

where $\mathbf{T}_s$ is the flow map of s hard spheres.
A popular method to establish the convergence of solutions of (1.1) is to first demonstrate that the mild form of the BBGKY hierarchy is given by a contracting operator if t is sufficiently small. This step ensures that the sum (1.2) converges absolutely, and in a second step one checks the convergence of the individual terms. It is noteworthy that the necessity to establish the contraction property of the mild BBGKY operator (and thereby restricting the analysis to small values of t) is due to the fact that the individual terms in (1.2) are unsigned.
The key achievement of this paper is to demonstrate that in a simpler setting a representation formula similar to (1.2) for $f_N^{(s)}$ can be found so that the individual terms are non-negative and can be interpreted as probabilities. This representation offers two significant advantages:

(1) It is possible to allow irregular initial data which may not be exponentially tight.

(2) The need to establish a contraction property is replaced by a tightness bound, which is related to the properties of physically relevant objects such as collision histories. In particular, convergence can be established for all times.

1.2. **The Lorentz and Rayleigh Gas.** Instead of considering a system of N identical hard spheres evolving via elastic collisions one can consider a single tagged or tracer particle evolving among a system of fluid scatterers or background particles.
If the background particles are fixed and of infinite relative mass to the tracer particle then one has a model known as the Lorentz gas first introduced by Lorentz in [21] to study the motion of electrons in a metal.
Much research has been done deriving the linear Boltzmann equation from a Lorentz gas with randomly placed scatterers, for example [7, 15, 30] and a large number of references found in Part I Chapter 8 of [32]. The linear Boltzmann equation can however fail to hold if we consider non-random periodic scatterers, as shown for example by Golse [16] and Marklof [22]. The existence of a limiting stochastic process for the periodic Lorentz gas from the Boltzmann-Grad limit was shown by Marklof and Strömbergsson in [23].
When a force field is present the convergence of the distribution of the tracer particle in an absorbing Lorentz gas to the solution of a gainless linear Boltzmann equation was proved in

[13]. The authors also proved that if the scatterers move with a constant random initial velocity then the convergence can be proven with significantly weaker assumptions on the force field.
Closely related to the Lorentz gas is the Rayleigh gas, where the background particles are no longer of infinite mass. Convergence to Brownian motion is discussed in Part I Chapter 8 of [32]. In [20] Lebowitz and Spohn proved the convergence of the momentum process for a test particle to a jump process associated to the linear Boltzmann equation. This was proved for arbitrarily long times, via the BBGKY hierarchy, when the initial distribution of the velocities is at equilibrium. This builds on their previous work [18, 19, 34].
In this paper we consider a Rayleigh gas where the background particles are of equal mass to the tagged particle and have no self interaction. The particles evolve via a simplified form of hard-sphere dynamics whereby the background particles do not change velocity. We consider non-equilibrium initial data but require that the background particles are spatially homogeneous. Convergence is proved for arbitrary times.

2. Model and Main Result

Our Rayleigh gas model in three dimensional space is now detailed. Define $U := [0,1]^3$ with periodic boundary conditions. Here a tagged particle evolves via the hard sphere flow and the remaining N particles do not interact, i.e. move along straight lines. The initial distribution of the tagged particle is $f_0 \in L^1(U \times \mathbb{R}^3)$. The N background particles are independently distributed according to the law $g_0 \in L^1(\mathbb{R}^3)$ in velocity space and uniform in U.
The tagged particle and the background particles are spheres with diameter $\varepsilon > 0$ which is related to N via the Boltzmann-Grad scaling,

$$N\varepsilon^2 = 1. \tag{2.1}$$

The background particles always travel in free flow with their velocities never changing from the initial value. The tagged particle travels in free flow whilst its centre remains at least ε away from the centre of all the background particles.
When the centre of the tagged particle comes within ε of the centre of a background particle the tagged particle collides as if it was a Newtonian hard-sphere collision and changes velocity. Explicitly this is described as follows. Denote the position and velocity of background particle $1 \le j \le N$ at time t by $(x_j(t), v_j(t))$. Then for all $t \ge 0$,

$$\frac{\mathrm{d}x_j(t)}{\mathrm{d}t} = v_j(t) \text{ and } \frac{\mathrm{d}v_j(t)}{\mathrm{d}t} = 0.$$

Further denote the position and velocity of the tagged particle at time t by $(x(t), v(t))$. Then for all $t \ge 0$,

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = v(t).$$

If at time t for all $1 \le j \le N$, $|x(t) - x_j(t)| > \varepsilon$ then $\mathrm{d}v(t)/\mathrm{d}t = 0$. Otherwise there exists a $1 \le j \le N$ such that $|x(t) - x_j(t)| = \varepsilon$ and the tagged particle experiences an instantaneous change of velocity. Define the collision parameter $\nu \in \mathbb{S}^2$ by,

$$\nu := \frac{x(t) - x_j(t)}{|x(t) - x_j(t)|}.$$

Then the velocity of the tagged particle instantaneously after the collision, $v(t)$, is given by

$$v(t) := v(t^-) - \nu \cdot (v(t^-) - v_j)\nu.$$

It is noted that in this model we do not have conservation of momentum. The background particles do not change velocity and the root particle does.

Proposition 2.1. *For $N \in \mathbb{N}$ and $T > 0$ fixed these dynamics are well defined up to time T for all initial configurations apart from a set of measure zero.*

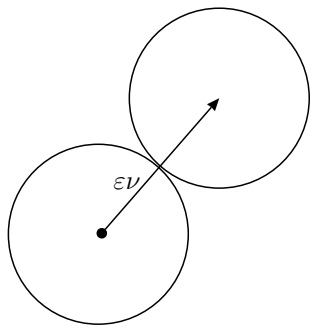


FIGURE 1. The collision parameter ν

The proof is given in section 6, which establishes that almost surely all collisions involve only pairs.
We are interested in studying the distribution of a tagged particle among N background particles, $\hat{f}_t^N$, under the above particle dynamics as N tends to infinity or equivalently as ε tends to zero.

Definition 2.2. *Probability densities $f_0 \in L^1(U \times \mathbb{R}^3)$, $g_0 \in L^1(\mathbb{R}^3)$ are admissible if*

$$\int_{U\times\mathbb{R}^3} f_0(x,v)(1+|v|^2)\,\mathrm{d}x\,\mathrm{d}v < \infty, \tag{2.2}$$

$$\int_{\mathbb{R}^3} g_0(v)(1+|v|^2)\,\mathrm{d}v < \infty, \tag{2.3}$$

$$\operatorname*{ess\,sup}_{v\in\mathbb{R}^3} g_0(v)(1+|v|^4) < \infty. \tag{2.4}$$

The distribution of the tagged particle is shown to converge to the solutions of a linear Boltzmann equation up to a finite arbitrary time T. We now state the main theorem of this paper.

Theorem 2.3. *Let $0 < T < \infty$ and f_0, g_0 be admissible. Then $\hat{f}_t^N$ converges to a time-dependent density f_t in the TV sense. Moreover, the limit f_t satisfies the linear Boltzmann equation*

$$\begin{cases} \partial_t f_t(x,v) &= -v\cdot\partial_x f_t(x,v) + Q[f_t](x,v), \\ f_{t=0}(x,v) &= f_0(x,v), \end{cases} \tag{2.5}$$

where the collision operator Q is defined by $Q := Q^+ - Q^-$ and Q^+ and Q^- known respectively as the gain and loss term are given as follows,

$$Q^+[f](x,v) = \int_{\mathbb{S}^2}\int_{\mathbb{R}^3} f(x,v')g_0(\bar{v}')[(v-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu,$$

where the pre-collision velocities, v' and $\bar{v}'$, are given by $v' = v+\nu\cdot(\bar{v}-v)\nu$ and $\bar{v}' = \bar{v}-\nu\cdot(\bar{v}-v)\nu$, and

$$Q^-[f](x,v) = f(x,v)\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu.$$

2.1. **Remarks.**

(1) The reader is reminded that solutions of (2.5) only conserve mass, but not energy.
(2) The analysis of the Rayleigh gas can also be done using traditional BBGKY approach. Here one uses the collision operator in $\mathbb{R}^d$:

$$\begin{aligned}\mathcal{C}^{\mathrm{Rayl}}_{s,s+1} f^{(s+1)}(t,Z_s) = (N-s)\varepsilon^{d-1}\int_{S^{d-1}\times\mathbb{R}^d} &\nu\cdot(v_{s+1}-v_i) \\ &\times f_N^{(s+1)}(t,Z_s,x_1+\varepsilon\nu,v_{s+1})\,\mathrm{d}\nu\,\mathrm{d}v_{s+1}.\end{aligned}$$

The hard sphere collision operator is given by

$$\mathcal{C}^{\mathrm{hs}}_{s,s+1} f^{(s+1)}(t, Z_s) := (N-s)\varepsilon^{d-1} \sum_{i=1}^{s} \int_{S^{d-1}\times\mathbb{R}^d} \nu \cdot (v_{s+1} - v_i) \times f_N^{(s+1)}(t, Z_s, x_i + \varepsilon\nu, v_{s+1})\, \mathrm{d}\nu\, \mathrm{d}v_{s+1}. \tag{2.6}$$

The only difference between $\mathcal{C}^{\mathrm{Rayl}}_{s,s+1}$ and $\mathcal{C}^{\mathrm{hs}}_{s,s+1}$ is the fact that in the hard sphere case one sums over all indices $i = 1, \dots, s$ and in the Rayleigh case only over $i = 1$. This gives estimates on $\mathcal{C}^{\mathrm{Rayl}}_{s,s+1}$, which are independent of s in the contraction proof for the mild form of the associated BBGKY hierarchy. Using the function spaces $\mathbf{X}_{\varepsilon,\beta,\mu}$ with norm $\|.\|_{\varepsilon,\beta,\mu}$ as in [14, Def 5.1.4] for measurable functions $G : t \in [0,T] \mapsto (t) = (g_s(t))_{s\geq 1} \in \mathbf{X}_{\varepsilon,\beta,\mu}$ one can introduce another time-dependent variant compared to [14, Def 5.1.4]

$$|||G|||_{\varepsilon,\beta,\mu,\lambda} := \sup_{0\leq t\leq T} \|G(t)\exp(-\lambda t|v_1|^2)\|_{\varepsilon,\beta,\mu}.$$

This will lead to a contraction for arbitrary large times T. For a slightly different approach assuming only finite moments see [31, Section II.B].

(3) Our method can be used to derive quantitative error estimates at the expense of more complex notation and additional regularity requirements for f_0 and g_0. In particular, see lemmas 4.14 and 4.17 for some quantitative expressions.

(4) The result should also hold in the case $d = 2$ or $d \geq 4$ up to a change in moment assumptions on the initial data and minor changes in estimates and calculations throughout the paper.

(5) One could consider a spatially inhomogeneous initial distribution for the background particles $g_0 = g_0(x, v)$. This adds a complication to the equations since for example the operator Q in (2.5) becomes time-dependent, i.e. $Q_t = Q_t^+ - Q_t^-$ with

$$Q_t^-[f](x,v) = f(x,v) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(x - t\bar{v}, \bar{v})[(v-\bar{v})\cdot\nu]_+ \,\mathrm{d}\bar{v}\,\mathrm{d}\nu,$$

and Q_t^+ analogous. Since the operator now depends on the time t this would require evolution semigroup results to echo the semigroup results in [3].

(6) One could also attempt to adapt these results to more complex and involved models. For example a model where each particle has an associated counter and a collision occurs between particle i and j if and only if both counters are less than k, in the hope of letting k tend to infinity. The main difficulty here will be that one will need to keep track of the current distribution of the background g_t in contrast to our model where the background has constant with time distribution g_0.

2.2. **Method of Proof.** We closely follow the method of [26]. That is we study the probability distribution of finding a given history of collisions, a given tree, at time t.
Firstly in section 3 we prove the main result, theorem 3.1, which shows that there exists a solution P_t to a Kolmogorov differential equation (3.1) and relate this solution to the solution of the linear Boltzmann equation. We show existence by explicitly building a solution on the most simple trees and using this to iteratively build a full solution.
In section 4 we consider the distribution $\hat{P}_t$ of finding a given history of collisions from our particle dynamics and show by direct calculation that this solves a similar differential equation in theorem 4.6 for sufficiently well controlled (good) trees.
Finally in section 5 we prove the main theorem of the paper, theorem 2.3, by proving the convergence between P_t and $\hat{P}_t$ in theorem 5.8 and then relating this to f_t and $\hat{f}_t^N$.

2.3. **Tree Set Up.** We construct trees in a similar way to [26].
A tree represents a specific history of collisions. The nodes of the tree are denoted by m and represent particles while the edges, denoted E, represent collisions. The root of the tree represents the tagged particle and is marked with the initial position of the tagged particle $(x_0, v_0) \in U \times \mathbb{R}^3$. The child nodes of the root represent background particles that the root collides with and are denoted $(t_j, \nu_j, v_j) \in (0, T] \times \mathbb{S}^2 \times \mathbb{R}^3$, where t_j represents the collision time, ν_j the collision parameter and v_j the incoming velocity of the background particle. Since the background particles only collide with the root particle and not each other we only consider trees of height at most 1, so the trees simplify to the initial position of the tagged particle and a list of its $n \geq 0$ collisions. The graph structure is mainly suppressed.

Definition 2.4. *The set of collision trees $\mathcal{MT}$ is defined by,*

$$\mathcal{MT} := \{(x_0, v_0), (t_1, \nu_1, v_1), \ldots, (t_n, \nu_n, v_n) : n \in \mathbb{N} \cup \{0\}\}.$$

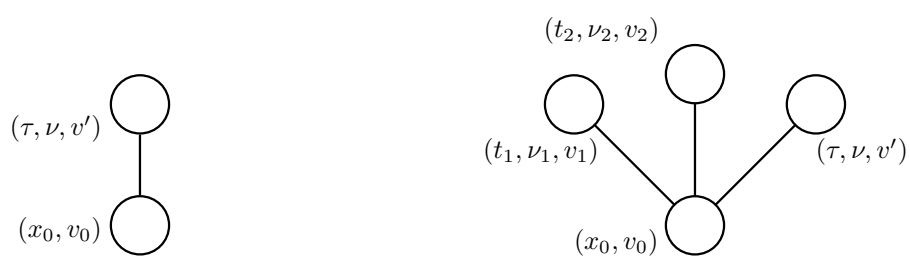
For a tree $\Phi \in \mathcal{MT}$, $n(\Phi)$ is the number of collisions.
The final collision in a tree Φ plays a significant role. Define the maximum collision time $\tau(\Phi) \in [0, T]$,

$$\tau(\Phi) := \begin{cases} 0 & \text{if } n(\Phi) = 0, \\ \max_{1 \leq j \leq n(\Phi)} t_j & \text{else.} \end{cases} \tag{2.7}$$

Further for $n(\Phi) \geq 1$ the marker for the final collision is denoted by,

$$(\tau, \nu, v') := (t_{n(\Phi)}, \nu_{n(\Phi)}, v_{n(\Phi)}).$$

The realisation of a tree Φ at a time $t \in [0, T]$ for a particle diameter $\varepsilon > 0$ uniquely defines the position and velocity of the tagged particle for all times up to t since the initial position and the collisions the root experiences are known. Further it determines the initial positions of the $n(\Phi)$ background particles involved in the tree since we can work backwards from the collision and we know that their velocity does not change. Finally it also includes information about the other $N - n$ background particles, because it is known that they do not interfere with the root up to time t.
If the root collides at the instant t denote the pre-collisional velocity by $v(t^-)$ and the post-collisional velocity by $v(t)$. Throughout this paper the dependence on Φ is often dropped from these and other variables when the context is clear.
For $n \geq 1$, define $\bar{\Phi}$ as being the pruned tree of Φ with the node representing the final collision, which occurs at time τ, removed. For example if $\Phi = ((x_0, v_0), (t_1, \nu_1, v_1), (\tau, \nu, v'))$ then $\bar{\Phi} = ((x_0, v_0), (t_1, \nu_1, v_1))$.



(a) An example tree with 1 collision

(b) An example tree with 3 collisions

FIGURE 2. Two example trees

Define a metric d on $\mathcal{MT}$ as follows. For $\Phi, \Psi \in \mathcal{MT}$ with components Φ_j and Ψ_j respectively.

$$\mathrm{d}(\Phi, \Psi) := \begin{cases} 1 & \text{if } n(\Phi) \neq n(\Psi) \\ \min\{1, \max_{0 \leq j \leq n} |\Phi_j - \Psi_j|_\infty\} & \text{else.} \end{cases}$$

Further denote by $B_h(\Phi)$ the ball of radius $h/2$ around $\Phi \in \mathcal{MT}$,

$$B_h(\Phi) := \{\Psi \in \mathcal{MT} : \mathrm{d}(\Phi, \Psi) < h/2\}. \tag{2.8}$$

The standard Lebesgue measure on $\mathcal{MT}$ is denoted by $\mathrm{d}\lambda$.

3. The Idealized Distribution

In this section we show that there exists a solution, denoted P_t, to the idealized equation, equation (3.1), and relate this solution to the solution of the linear Boltzmann equation. We prove existence by constructing a solution iteratively on different sized trees. In section 5 P_t is compared to the solution of a similar evolution equation defined by the particle dynamics in order to show the required convergence. The idealized system plays the same role as the Boltzmann hierarchy in [14].

For a given tree $\Phi \in \mathcal{MT}$, $P_0(\Phi)$ is zero unless Φ involves no collisions, in which case $P_0(\Phi)$ is given in terms of the initial distribution f_0. $P_t(\Phi)$ remains zero until $t = \tau$ when there is an instantaneous increase to a positive value depending on $P_\tau(\bar{\Phi})$ and the final collision in Φ. For $t > \tau$, $P_t(\Phi)$ decreases at a rate that is obtained by considering all possible collisions.

The idealized equation is given by,

$$\begin{cases} \partial_t P_t(\Phi) &= \mathcal{Q}_t[P_t](\Phi) = \mathcal{Q}_t^+[P_t](\Phi) - \mathcal{Q}_t^-[P_t](\Phi), \\ P_0(\Phi) &= f_0(x_0, v_0)\mathbb{1}_{n(\Phi)=0}, \end{cases} \tag{3.1}$$

where,

$$\mathcal{Q}_t^+[P_t](\Phi) := \begin{cases} \delta(t-\tau)P_t(\bar{\Phi})g_0(v')[(v(\tau^-) - v') \cdot \nu]_+ & \text{if } n(\Phi) \geq 1, \\ 0 & \text{if } n(\Phi) = 0, \end{cases} \tag{3.2}$$

$$\mathcal{Q}_t^-[P_t](\Phi) := P_t(\Phi) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau) - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu. \tag{3.3}$$

Theorem 3.1. *Suppose that f_0 and g_0 are admissible (in the sense of Def. 2.2). Then there exists a solution $P : [0, T] \to L^1(\mathcal{MT})$ to the idealized equation,* (3.1)*. Moreover for any $t \in [0, T]$ and for any $\Omega \subset U \times \mathbb{R}^3$ define*

$$S_t(\Omega) := \{\Phi \in \mathcal{MT} : (x(t), v(t)) \in \Omega\}.$$

Then

$$\int_\Omega f_t(x, v) \, \mathrm{d}x \, \mathrm{d}v = \int_{S_t(\Omega)} P_t(\Phi) \, \mathrm{d}\Phi,$$

where f_t is the unique mild solution of the linear Boltzmann equation given in proposition 3.5.

Remark 3.2. *Condition* (2.4) *can be relaxed to*

$$\operatorname*{ess\,sup}_{v \in \mathbb{R}^3} g_0(v)(1 + |v|^{3+\eta}) < \infty \tag{3.4}$$

for some $\eta > 0$.

From now on assume that f_0 and g_0 are admissible with the provision that either (2.4) or (3.4) holds. We prove the existence by construction, taking several steps to build a solution by solving on the most simple trees first and using this solution to iteratively build a full solution. We begin by solving the linear Boltzmann equation. We establish existence, uniqueness and regularity of solutions of (2.5) by adapting methods from semigroup theory. The difficulty here is that after writing the linear Boltzmann equation as the sum of two unbounded operators we need to ensure that a honest semigroup is generated in order to prove existence and uniqueness. Next we adapt these semigroup techniques to define functions $P_t^{(j)}$ that describe the distribution of finding the tagged particle such that it has experienced j collisions. This is key to connecting P_t to the solution of the linear Boltzmann.

The following notion of mild solution suitable for transport equations is used (c.f. [1, Def 3.1.1])

Definition 3.3. *Consider the following system,*

$$\begin{cases} \partial_t u(t) &= Lu(t), \\ u(0) &= u_0. \end{cases} \tag{3.5}$$

Where $L : D(L) \subset L^1(U \times \mathbb{R}^3) \to L^1(U \times \mathbb{R}^3)$ is an operator and $u_0 \in L^1(U \times \mathbb{R}^3)$ is given. The function $u : [0, T] \to U \times \mathbb{R}^3$ is called a mild solution of (3.5) if for all $t \geq 0$,

$$\int_0^t u(\theta)\,\mathrm{d}\theta \in D(L) \;\text{ and }\; L\int_0^t u(\theta)\,\mathrm{d}\theta = u(t) - u_0.$$

We split the right hand side of (2.5) into two operators, A and B. These will appear in the construction of P_t.

Definition 3.4. *Define $D(A), D(B) \subset L^1(U \times \mathbb{R}^3)$ by,*

$$D(A) := \{f \in L^1(U \times \mathbb{R}^3) : v \cdot \partial_x f(x, v) + Q^-[f](x, v) \in L^1(U \times \mathbb{R}^3)\},$$
$$D(B) := \{f \in L^1(U \times \mathbb{R}^3) : Q^+[f] \in L^1(U \times \mathbb{R}^3)\}.$$

Then define $A : D(A) \to L^1(U \times \mathbb{R}^3)$ and $B : D(B) \to L^1(U \times \mathbb{R}^3)$ by,

$$(Af)(x, v) := -v \cdot \partial_x f(x, v) - Q^-[f](x, v), \tag{3.6}$$

$$(Bf)(x, v) := Q^+[f](x, v). \tag{3.7}$$

Proposition 3.5. *Suppose that the assumptions in theorem 3.1 hold. Then there exists a unique mild solution $f : [0, T] \to L^1(U \times \mathbb{R}^3)$ to (2.5). Furthermore f_t remains non-negative and of mass 1, and*

$$\int_{U \times \mathbb{R}^3} f_t(x, v)(1 + |v|)\,\mathrm{d}x\,\mathrm{d}v < \infty, \tag{3.8}$$

$$f_t \in D(B) \tag{3.9}$$

hold for all $t \in [0, T]$.

Proof. See section 6. □

Proposition 3.6. *There exists a unique mild solution, $P^{(0)} : [0, T] \to L^1(U \times \mathbb{R}^3)$, to the following evolution equation,*

$$\begin{cases} \partial_t P_t^{(0)}(x, v) &= (AP_t^{(0)})(x, v), \\ P_0^{(0)}(x, v) &= f_0(x, v). \end{cases} \tag{3.10}$$

Where A is as in (3.6).

The distribution $P_t^{(0)}(x, v)$ can be thought of as the probability of finding the tagged particle at (x, v) at time t such that it has not yet experienced any collisions.

Proof. By lemma 6.1 A generates the substochastic C_0 semigroup $T(t)$ given in (6.1). By the Hille-Yoshida theorem, [28, Thm 1.3.1] A is closed. By [1, Thm 3.1.12] (3.10) has a unique mild solution given by

$$P_t^{(0)} = T(t)f_0. \tag{3.11}$$

□

Lemma 3.7. *For all $t \in [0, T]$, $P_t^{(0)} \leq f_t$ pointwise.*

Remark 3.8. *This lemma is entirely expected. The probability of finding the tagged particle at (x, v) at time t is given by $f_t(x, v)$ and the probability of finding it at (x, v) at time t such that it has not experienced any collisions up to time t is given by $P_t^{(0)}(x, v)$ so one expects $P_t^{(0)} \leq f_t$.*

Proof. For $t \in [0,T]$ define $F_t^{(0)} := f_t - P_t^{(0)}$. Then since f_t and $P_t^{(0)}$ are mild solutions of (2.5) and (3.10) respectively, $F_t^{(0)}$ is a mild solution of

$$\begin{cases} \partial_t F_t^{(0)}(x,v) &= AF_t^{(0)} + Bf_t(x,v) \\ F_0^{(0)}(x,v) &= 0. \end{cases}$$

By (3.9) and [1, Prop. 3.1.16] $F_t^{(0)}$ is given by,

$$F_t^{(0)} = \int_0^t T(t-\theta)Bf_\theta \,\mathrm{d}\theta.$$

Now noting that f_θ is non-negative it follows that Bf_θ and hence $T(t-\theta)Bf_\theta$ are non-negative also. Hence $F_t^{(0)} \geq 0$ which implies $P_t^{(0)} \leq f_t$. □

Definition 3.9. *For $j \in \mathbb{N} \cup \{0\}$ denote by $\mathcal{T}_j$ the set of all trees with exactly j collisions. Explicitly,*

$$\mathcal{T}_j := \{\Phi \in \mathcal{MT} : n(\Phi) = j\}. \tag{3.12}$$

The required solution P_t can now be defined iteratively on the space $\mathcal{MT}$. For $\Phi \in \mathcal{T}_0$ define

$$P_t(\Phi) := P_t^{(0)}(x(t), v(t)). \tag{3.13}$$

Else define,

$$\begin{aligned} P_t(\Phi) := \mathbb{1}_{t\geq\tau} \exp\left(-(t-\tau)\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu']_+ \,\mathrm{d}\bar{v}\,\mathrm{d}\nu'\right) \\ P_\tau(\bar{\Phi})g_0(v')[(v(\tau^-)-v')\cdot\nu]_+. \end{aligned} \tag{3.14}$$

The right hand side of this equation depends on $P_\tau(\bar{\Phi})$ but since $\bar{\Phi}$ has degree exactly one less than Φ the equation is well defined.

The proof that P_t has the required properties of theorem 3.1 is given shortly. We first define the function $P_t^{(j)}$ which is thought of, in parallel to $P_t^{(0)}$, as being the probability of finding the tagged particle at a certain position at time t such that it has experienced exactly j collisions up to time t. The $P_t^{(j)}$ will be required to show the connection between P_t and the solution of the linear Boltzmann equation.

Definition 3.10. *Let $t \in [0,T]$ and $\Omega \subset U \times \mathbb{R}^3$ be measurable. Recall in theorem 3.1 we define the set, $S_t(\Omega) = \{\Phi \in \mathcal{MT} : (x(t), v(t)) \in \Omega\}$ - the set of all trees such that the tagged particle at time t is in Ω. Define for all $j \in \mathbb{N} \cup \{0\}$,*

$$S_t^j(\Omega) := \mathcal{T}_j \cap S_t(\Omega).$$

Then for $j \geq 1$, define $P^{(j)}(\Omega)$ by,

$$P_t^{(j)}(\Omega) := \int_{S_t^j(\Omega)} P_t(\Phi)\,\mathrm{d}\Phi.$$

Lemma 3.11. *Let $t \in [0,T]$, $j \geq 1$. Then $P_t^{(j)}$ is absolutely continuous with respect to the Lebesgue measure on $U \times \mathbb{R}^3$.*

Proof. Let $j = 1$. Then we have by (3.14),

$$\begin{aligned} P_t^{(1)}(\Omega) &= \int_{S_t^1(\Omega)} P_t(\Phi) \\ &= \int_0^t \int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_U \exp\left(-(t-\tau)\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+ \,\mathrm{d}\bar{v}\,\mathrm{d}\nu\right) \\ &\qquad P_\tau^{(0)}(x_0+\tau v_0, v_0)g_0(v')[(v_0-v')\cdot\nu]_+ \mathbb{1}_{(x(t),v(t))\in\Omega)}\,\mathrm{d}x_0\,\mathrm{d}\nu\,\mathrm{d}v'\,\mathrm{d}v_0\,\mathrm{d}\tau \end{aligned} \tag{3.15}$$

We define a coordinate transform $(\nu, x_0, v_0, v') \mapsto (\nu, x, v, \bar{w})$ given by,

$$\begin{aligned} v &:= v_0 + \nu(v' - v_0) \cdot \nu \\ x &:= x_0 + \tau v_0 + (t - \tau) v \\ \bar{w} &:= v' - \nu(v' - v_0) \cdot \nu. \end{aligned}$$

This transformation has Jacobi matrix,

$$\begin{pmatrix} \mathrm{Id} & 0 & 0 & 0 \\ & \mathrm{Id} & & \\ & 0 & \mathrm{Id} - \nu \otimes \nu & \nu \otimes \nu \\ & 0 & \nu \otimes \nu & \mathrm{Id} - \nu \otimes \nu \end{pmatrix}$$

where the blank entries are not required for the computation of the matrix's determinant. The 2x2 matrix in the bottom right has determinant -1 and hence the absolute value of the determinant of the entire matrix is 1. With this (3.15) becomes,

$$\begin{aligned} P_t^{(1)}(\Omega) = \int_\Omega \int_0^t \int_{\mathbb{R}^3} \int_{\mathbb{S}^2} \exp\left(-(t-\tau) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu \right) \\ P_\tau^{(0)}(x - (t-\tau)v, w') g_0(\bar{w}')[(v - \bar{w}) \cdot \nu]_+ \, \mathrm{d}\nu \, \mathrm{d}\bar{w} \, \mathrm{d}\tau \, \mathrm{d}x \, \mathrm{d}v, \end{aligned}$$

where $w' = v + \nu(\bar{w} - v) \cdot \nu$ and $\bar{w}' = \bar{w} - \nu(\bar{w} - v) \cdot \nu$. Hence we see that if the Lebesgue measure of Ω equals zero then so does $P_t^{(1)}(\Omega)$. For $j \geq 2$ we use a similar approach, using the iterative formula for $P_t(\Phi)$ (3.14). □

Remark 3.12. *Since $P_t^{(j)}$ is an absolutely continuous measure on $U \times \mathbb{R}^3$, the Radon-Nikodym theorem (see Theorem 4.2.2* [12]*) implies that $P_t^{(j)}$ has a density, which we denote by $P_t^{(j)}$ also. This gives,*

$$\int_\Omega P_t^{(j)}(x, v) \, \mathrm{d}x \, \mathrm{d}v = \int_{S_t^j(\Omega)} P_t(\Phi) \, \mathrm{d}\Phi.$$

Hence for almost all $(x, v) \in U \times \mathbb{R}^3$,

$$P_t^{(j)}(x, v) = \int_{S_t^j(x,v)} P_t(\Phi) \, \mathrm{d}\Phi.$$

Remark 3.13. *A similar formula holds for $P_t^{(0)}$ since the set $S_t^0(x, v)$ contains exactly one tree: the tree Φ with initial root data $(x - tv, v)$ and such that the root has no collisions,*

$$\int_{S_t^0(x,v)} P_t(\Phi) \, \mathrm{d}\Phi = P_t((x - tv, v)) = P_t^{(0)}(x, v).$$

Proposition 3.14. *For $j \geq 1$, $P_t^{(j)}$ as defined above is almost everywhere the unique mild solution to the following differential equation,*

$$\begin{cases} \partial_t P_t^{(j)}(x, v) & = A P_t^{(j)}(x, v) + B P_t^{(j-1)}(x, v), \\ P_0^{(j)}(x, v) & = 0, \end{cases}$$

where A is given in (3.6) *and B in* (3.7)*.*

The following lemma helps prove the proposition for the case $j = 1$ which allows the use of an inductive argument to prove the proposition in full.

Lemma 3.15. *For any $t \in [0, T]$ and almost all $(x, v) \in U \times \mathbb{R}^3$,*

$$P_t^{(1)}(x, v) = \int_0^t T(t - \theta) B P_\theta^{(0)}(x, v) \, \mathrm{d}\theta,$$

where the semigroup $T(t)$ is as in (6.1)*. The right hand side is well defined since* (3.9)*,* (3.7) *and* (6.8) *imply $P_t^{(0)} \in D(B)$.*

Proof. We show that for any $\Omega \subset U \times \mathbb{R}^3$ measurable we have,

$$\int_\Omega P_t^{(1)}(x,v)\,\mathrm{d}x\,\mathrm{d}v = \int_\Omega \int_0^t T(t-\theta)BP_\theta^{(0)}(x,v)\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}v.$$

By the definition of $T(t)$ in equation (6.1), the definition of B in (3.7), and the proof of lemma 3.11, for $w' = v + \nu(\bar{w} - v)\cdot\nu$ and $\bar{w}' = \bar{w} - \nu(\bar{w} - v)\cdot\nu$, we have,

$$\begin{aligned}
&\int_\Omega \int_0^t T(t-\theta)BP_\theta^{(0)}(x,v)\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}v \\
&\quad= \int_\Omega \int_0^t \exp\left(-(t-\theta)\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu\right) \\
&\qquad\qquad \times BP_\theta^{(0)}(x-(t-\theta)v,v)\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}v \\
&\quad= \int_\Omega \int_0^t \exp\left(-(t-\theta)\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu\right) \\
&\qquad\quad \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} P_\theta^{(0)}(x-(t-\theta)v,w')g_0(\bar{w}')[(v-\bar{w})\cdot\nu]_+\,\mathrm{d}\nu\,\mathrm{d}\bar{w}\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}v \\
&\quad= \int_\Omega \int_0^t \int_{\mathbb{R}^3}\int_{\mathbb{S}^2} \exp\left(-(t-\theta)\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu\right) \\
&\qquad\qquad P_\theta^{(0)}(x-(t-\theta)v,w')g_0(\bar{w}')[(v-\bar{w})\cdot\nu]_+\,\mathrm{d}\nu\,\mathrm{d}\bar{w}\,\mathrm{d}\theta\,\mathrm{d}x\,\mathrm{d}v \\
&\quad= \int_\Omega P_t^{(1)}(x,v)\,\mathrm{d}x\,\mathrm{d}v.
\end{aligned}$$

□

Proof of proposition 3.14. Consider induction on j. First let $j = 1$. We seek to apply [1, Prop 3.1.16]. If $\int_0^t BP_\theta^{(0)}\,\mathrm{d}\theta \in L^1(U \times \mathbb{R}^3)$ then the proposition holds so by the above lemma $P_t^{(1)}$ is the unique mild solution.
To this aim note that since $P_t^{(0)}$ is the unique mild solution of (3.10),

$$\int_0^t P_\theta^{(0)}\,\mathrm{d}\theta \in D(A).$$

By [3, Section 10.4.3] $D(A) \subset D(B)$ and hence

$$\int_0^t P_\theta^{(0)}\,\mathrm{d}\theta \in D(B).$$

This implies,

$$B\int_0^t P_\theta^{(0)}\,\mathrm{d}\theta \in L^1(U \times \mathbb{R}^3).$$

as required. Now consider $j \geq 2$ and assume the proposition is true for $j-1$. By setting $F_t^{(j-1)} := f_t - P_t^{(j-1)}$ a similar argument to lemma 3.7 shows that $P_t^{(j-1)} \leq f_t$. By (3.9) and (6.8), $P_t^{(j-1)} \in D(B)$ so the right hand side is well defined. A similar approach to lemma 3.15 shows that for any $t \in [0,T]$ and almost all $(x,v) \in U \times \mathbb{R}^3$,

$$P_t^{(j)}(x,v) = \int_0^t (T(t-\theta)BP_\theta^{(j-1)})(x,v)\,\mathrm{d}\theta,$$

where $T(t)$ is the semigroup given in (6.1). The rest follows by the same argument as in the $j = 1$ case. □

Proposition 3.16. *For all $t \in [0,T]$ and almost all $(x,v) \in U \times \mathbb{R}^3$,*

$$\sum_{j=0}^{\infty} P_t^{(j)}(x,v) = f_t(x,v), \tag{3.16}$$

where f_t is the unique mild solution of the linear Boltzmann equation given in proposition 3.5.

Proof. Since $P_t^{(0)}$ is a mild solution of (3.10),

$$\int_0^t P_\theta^{(0)}(x,v)\,\mathrm{d}\theta \in D(A), \tag{3.17}$$

and,

$$P_t^{(0)}(x,v) = f_0(x,v) + A\int_0^t P_\theta^{(0)}(x,v)\,\mathrm{d}\theta. \tag{3.18}$$

Further for $j \geq 1$ by proposition 3.14 and [3, Prop 3.31],

$$\int_0^t P_\theta^{(j)}(x,v)\,\mathrm{d}\theta \in D(A), \tag{3.19}$$

and,

$$P_t^{(j)}(x,v) = A\int_0^t P_\theta^{(j)}(x,v)\,\mathrm{d}\theta + \int_0^t BP_\theta^{(j-1)}(x,v)\,\mathrm{d}\theta. \tag{3.20}$$

Combining (3.17) and (3.19),

$$\int_0^t \sum_{j=0}^\infty P_\theta^{(j)}(x,v)\,\mathrm{d}\theta \in D(A).$$

Recalling from the proof of proposition 3.14 that $D(A) \subset D(B)$,

$$\int_0^t \sum_{j=0}^\infty P_\theta^{(j)}(x,v)\,\mathrm{d}\theta \in D(A)\cap D(B) = D(A+B).$$

Further summing (3.18) and (3.20) for $j \geq 1$,

$$\begin{aligned}\sum_{j=0}^\infty P_t^{(j)}(x,v) &= f_0(x,v) + \sum_{j=0}^\infty A\int_0^t P_\theta^{(j)}(x,v)\,\mathrm{d}\theta + \sum_{j=1}^\infty \int_0^t BP_\theta^{(j-1)}(x,v)\,\mathrm{d}\theta \\ &= f_0(x,v) + A\int_0^t \sum_{j=0}^\infty P_\theta^{(j)}(x,v)\,\mathrm{d}\theta + \int_0^t B\sum_{j=0}^\infty P_\theta^{(j)}(x,v)\,\mathrm{d}\theta \\ &= f_0(x,v) + (A+B)\int_0^t \sum_{j=0}^\infty P_\theta^{(j)}(x,v).\end{aligned}$$

Hence by definition 3.3, $\sum_{j=0}^\infty P_t^{(j)}(x,v)$ is a mild solution of (6.6) and therefore since f_t is the unique mild solution the proof is complete. □

We now have all the results needed to prove that P_t satisfies all the requirements of theorem 3.1.

Proof of theorem 3.1. Using definition 3.10, proposition 3.16, and, since each $P_t^{(j)}$ is positive, the monotone convergence theorem we have for any measurable $\Omega \subset U \times \mathbb{R}^3$,

$$\begin{aligned}\int_{S_t(\Omega)} P_t(\Phi)\,\mathrm{d}\Phi &= \sum_{j=0}^\infty \int_{S_t^j(\Omega)} P_t(\Phi)\,\mathrm{d}\Phi = \sum_{j=0}^\infty \int_\Omega P_t^{(j)}(x,v)\,\mathrm{d}x\,\mathrm{d}v \\ &= \int_\Omega f_t(x,v)\,\mathrm{d}x\,\mathrm{d}v.\end{aligned} \tag{3.21}$$

In particular,

$$\int_{\mathcal{MT}} P_t(\Phi)\,\mathrm{d}\Phi = \int_{U\times\mathbb{R}^3} f_t(x,v)\,\mathrm{d}x\,\mathrm{d}v < \infty.$$

Hence $P_t \in L^1(\mathcal{MT})$. To show that P_t is a solution of (3.1) first consider $\Phi \in \mathcal{T}_0$. Since $n(\Phi) = 0$,

$$P_0(\Phi) = P_0^{(0)}(x(0), v(0)) = f_0(x_0, v_0) = f_0(x_0, v_0)\mathbb{1}_{n(\Phi)=0}.$$

Hence it solves the initial condition. Now for $t > 0$, since $\Phi \in \mathcal{T}_0$, $v(t) = v_0$ and $x(t) = x_0 + tv_0$. Hence by (6.1) and (3.11),

$$\begin{aligned} P_t(\Phi) &= P_t^{(0)}(x(t), v(t)) = P_t^{(0)}(x_0 + tv_0, v_0) \\ &= \exp\left(-t \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v_0 - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu\right) f_0(x_0, v_0). \end{aligned}$$

The only dependence on t here is in the exponential term so we differentiate $P_t(\Phi)$ with respect to t,

$$\begin{aligned} \partial_t P_t(\Phi) &= \partial_t \left(\exp\left(-t \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v_0 - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu\right) f_0(x_0, v_0)\right) \\ &= -\int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v_0 - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu \times P_t(\Phi) = -\mathcal{Q}_t^-[P_t](\Phi). \end{aligned}$$

Hence P_t solves (3.1) on $\mathcal{T}_0$.
We now consider $\Phi \in \mathcal{T}_j$ for $j \geq 1$. Since $\Phi \in \mathcal{T}_j$ we have $n(\Phi) = j$ and $\tau > 0$. Hence

$$P_0(\Phi) = 0 = f_0(x_0, v_0)\mathbb{1}_{n(\Phi)=0}. \tag{3.22}$$

For $t = \tau$,

$$P_\tau(\Phi) = P_\tau(\bar{\Phi}) g_0(v')[v(\tau^-)_- v') \cdot \nu]_+. \tag{3.23}$$

Further for $t > \tau$ the only dependence on t is inside the exponential term and hence differentiating gives,

$$\begin{aligned} \partial_t P_t(\Phi) &= \partial_t \Big(\exp\big(-(t-\tau) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau) - \bar{v}) \cdot \nu']_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu'\big) \\ &\qquad P_\tau(\bar{\Phi}) g_0(v')[(v_0 - v') \cdot \nu]_+ \Big) \\ &= \exp\left(-(t-\tau) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau) - \bar{v}) \cdot \nu']_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu'\right) \\ &\qquad P_\tau(\bar{\Phi}) g_0(v')[(v_0 - v') \cdot \nu]_+ \\ &\qquad \times \left(-\int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau) - \bar{v}) \cdot \nu']_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu'\right) \\ &= -P_t(\Phi) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau) - \bar{v}) \cdot \nu']_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu' = -\mathcal{Q}_t^-[P_t](\Phi). \end{aligned} \tag{3.24}$$

Equations (3.22), (3.23) and (3.24) prove that P_t solves (3.1) on $\mathcal{T}_j$. Since $\mathcal{MT}$ is the disjoint union of $\mathcal{T}_j$ for $j \geq 0$, P_t is a solution of (3.1) on $\mathcal{MT}$. Finally, the required connection between P_t and the solution of the linear Boltzmann equation has been shown in (3.21). □

4. The Empirical Distribution

We now consider the empirical distribution of trees $\hat{P}_t^\varepsilon$ defined by the dynamics of the particle system for particles with diameter ε. To ease notation we drop the dependence on ε and write $\hat{P}_t$. The key result of this section is that $\hat{P}_t$ solves the differential equation (4.2) which is similar to the idealized equation (3.1). The similarity between the two equations is exploited in the next section to prove the required convergence as ε tends to zero.
We do this by restricting our attention to trees that are well controlled in various ways, calling these trees good trees.

Definition 4.1. *For a tree $\Phi \in \mathcal{MT}$ define $\mathcal{V}(\Phi) \in [0,\infty)$ to be the maximum velocity involved in the tree. That is,*

$$\mathcal{V}(\Phi) := \max\left\{\max_{j=1,\dots,n(\Phi)}\{|v_j|\},\ \max_{s\in[0,T]}\{|v(s)|\}\right\}.$$

Definition 4.2. *A tree $\Phi \in \mathcal{MT}$ is called re-collision free at diameter ε if for all $0 \le \varepsilon' \le \varepsilon$, for all $1 \le j \le n(\Phi)$ and for all $t < t_j$,*

$$|x(t) - (x_j + tv_j)| > \varepsilon'.$$

That is to say, if the tree involves a collision between the root particle and background particle j at time t_j then the root particle has not previously collided with background particle j. So if a tree is re-collision free then it involves at most one collision per background particle.
Further define

$$R(\varepsilon) := \{\Phi \in \mathcal{MT} : \Phi \text{ is re-collision free at diameter } \varepsilon\}.$$

Definition 4.3. *A tree $\Phi \in \mathcal{MT}$ is called non-grazing if all collisions in Φ are non-grazing, that is if,*

$$\min_{1\le j\le n(\Phi)} \nu_j \cdot (v(t_j^-) - v_j) > 0.$$

Definition 4.4. *A tree is $\Phi \in \mathcal{MT}$ is called free from initial overlap at diameter $\varepsilon > 0$ if initially the root is at least ε away from the centre of each background particle. Explicitly if, for $j = 1,\dots,N$,*

$$|x_0 - x_j| > \varepsilon.$$

Define $S(\varepsilon) \subset \mathcal{MT}$ to be the set of all trees that are free from initial overlap at radius ε.

Definition 4.5. *For any pair of decreasing functions $V, M : (0,\infty) \to [0,\infty)$ such that $\lim_{\varepsilon\to 0} V(\varepsilon) = \lim_{\varepsilon\to 0} M(\varepsilon) = \infty$, the set of good trees of diameter ε, $\mathcal{G}(\varepsilon)$, is defined as,*

$$\mathcal{G}(\varepsilon) := \Big\{\Phi \in \mathcal{MT} : n(\Phi) \le M(\varepsilon),\ \mathcal{V}(\Phi) \le V(\varepsilon),$$
$$\Phi \in R(\varepsilon) \cap S(\varepsilon) \text{ and } \Phi \text{ is non-grazing}\Big\}$$

Since M, V are decreasing for $\varepsilon' < \varepsilon$ we have $\mathcal{G}(\varepsilon) \subset \mathcal{G}(\varepsilon')$. Later some conditions on M and V are required to prove that $\hat{P}_t$ solves the relevant equation and to prove convergence.

Now define the operator $\hat{\mathcal{Q}}_t$ which mirrors the idealized operator $\mathcal{Q}_t$ in the empirical case. Fix $\hat{C}_1 > 0$, a constant depending only on Φ described later. Define the gain operator,

$$\hat{\mathcal{Q}}_t^+[\hat{P}_t](\Phi) := \begin{cases} \delta(t-\tau)\hat{P}_t(\bar{\Phi})\frac{g_0(v')[(v(\tau^-)-v')\cdot\nu]_+}{1-\pi\varepsilon^2\int_0^\tau |v(s)-v'|\,\mathrm{d}s+\hat{C}_1\varepsilon^3} & \text{if } n \ge 1 \\ 0 & \text{if } n = 0. \end{cases}$$

Next for a given tree Φ, a time $0 < t < T$ and $\varepsilon > 0$, define the function $\mathbb{1}_t^\varepsilon[\Phi] : U \times \mathbb{R}^3 \to \{0,1\}$ by

$$\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v}) := \begin{cases} 1 \text{ if for all } s \in (0,t),\ |x(s) - (\bar{x} + s\bar{v})| > \varepsilon, \\ 0 \text{ else.} \end{cases} \tag{4.1}$$

That is $\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})$ is 1 if a background particle starting at the position $(\bar{x},\bar{v})$ avoids colliding with the root particle of the tree Φ up to the time t. This allows us to define the loss operator,

$$\hat{\mathcal{Q}}_t^-[\hat{P}_t](\Phi) := \hat{P}_t(\Phi)\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu - \hat{C}_2(\varepsilon)}{\int_{U\times R^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}.$$

For some $\hat{C}_2(\varepsilon) > 0$ depending on t and Φ of $o(1)$ as ε tends to zero. Finally define the operator $\hat{\mathcal{Q}}_t$ as follows,

$$\hat{\mathcal{Q}}_t = \hat{\mathcal{Q}}_t^+ - \hat{\mathcal{Q}}_t^-.$$

Theorem 4.6. *For ε sufficiently small and for $\Phi \in \mathcal{G}(\varepsilon)$, $\hat{P}_t$ solves the following*

$$\begin{cases} \partial_t \hat{P}_t(\Phi) & = (1-\gamma(t))\hat{\mathcal{Q}}_t[\hat{P}_t](\Phi) \\ \hat{P}_0(\Phi) & = \zeta(\varepsilon) f_0(x_0, v_0)\mathbb{1}_{n(\Phi)=0}. \end{cases} \tag{4.2}$$

The functions γ and ζ are given by

$$\zeta(\varepsilon) := (1 - \frac{4}{3}\pi\varepsilon^3)^N, \tag{4.3}$$

and,

$$\gamma(t) := \begin{cases} n(\bar{\Phi})\varepsilon^2 & \text{if } t = \tau, \\ n(\Phi)\varepsilon^2 & \text{if } t > \tau. \end{cases}$$

Remark 4.7. *When choosing the background particles according to some Poisson point process some of these terms simplify as in* [25].

The proof is developed by a series of lemmas in which we prove the gain term, loss term and initial condition separately.

Definition 4.8. *Define $\omega_0 := (u_0, w_0) \in U \times \mathbb{R}^3$ to be the random initial position of the test particle. By our model ω_0 has distribution f_0.*
Further for $j = 1, \ldots, N$ define $\omega_j := (u_j, w_j)$ to be the random initial position and velocity of background particle j. Note that ω_j has distribution Unif$(U) \times g_0$. Finally define $\omega := (\omega_1, \ldots, \omega_N)$.

Lemma 4.9. *Let $\varepsilon > 0$ and $\Psi \in \mathcal{G}(\varepsilon)$ then $\hat{P}_t$ is absolutely continuous with respect to the Lebesgue measure λ on a neighbourhood of Ψ.*

Proof. Recall the definition of $B_h(\Psi)$ (2.8). Since $\mathcal{G}(\varepsilon)$ is open there exists a $h > 0$ such that $B_h(\Psi) \subset \mathcal{G}(\varepsilon)$. In the case $n(\Psi) = 0$ for all $t \geq 0$, $\hat{P}_t(\Psi) \leq f_0(x_0, v_0)$ and hence absolute continuity follows.
Suppose $n(\Phi) \geq 1$. Define a map $\varphi : B_h(\Psi) \to \mathcal{MT} \times U \times \mathbb{R}^3$,

$$\varphi(\Phi) := (\bar{\Phi}, (x(\tau) + \varepsilon\nu - \tau v', v')).$$

We view φ as having $n(\Phi) + 1$ components, the first being the initial root position (x_0, v_0), components $j = 2 \ldots, n$ being the marker (t_j, ν_j, ν_j) and the final component being $(x(\tau) + \varepsilon\nu - \tau v', v')$ - the initial position of the background particle that leads to the final collision with the root in Φ. We claim that,

$$\det(\nabla\varphi)(\Phi) = \varepsilon^2(v(\tau^-) - v') \cdot \nu. \tag{4.4}$$

To prove this we first rotate our coordinate axis so that $\nu = \mathbf{e}_1$. Then for $k = 0, \ldots, n$ define $F_{0,k} := \nabla_{x_0}\varphi_k(\Phi)$ and for $j = 1, \ldots, n$ define $F_{j,k} := \nabla_{t_k,\nu_k}\varphi_j(\Phi)$. We calculate,

$$F_{n+1,n+1} = \nabla_{\tau,\nu}\varphi_{n+1}(\Phi) = \begin{pmatrix} (v(\tau^-) - v') \cdot \nu & 0 & 0 \\ & \varepsilon & 0 \\ & 0 & \varepsilon \end{pmatrix},$$

where the blank components are not needed. Also, $F_{0,n} = \nabla_{x_0}\varphi_n(\Phi) = \mathrm{Id}(2)$. Further for $j = 2, \ldots, n+1$, $F_{j,j} = \nabla_{t_j,\nu_j}\varphi_j(\Phi) = \mathrm{Id}(2)$. For any other j, k not already calculated, $F_{j,k} = 0$. Hence $\det(\nabla\varphi)(\Phi)$ is the product of the determinants of all $F_{k,k}$ for $k = 1, \ldots, n+1$, proving the claim.
Now define a second map, $\tilde{\varphi} : B_h(\Psi) \to (U \times \mathbb{R})^{n+1}$,

$$\tilde{\varphi}(\Phi) := ((x_0, v_0), (x_1, v_1), \ldots, (x_n, v_n)).$$

This maps Φ to the initial position of each particle in Φ. By repeatedly applying (4.4),

$$\det(\nabla\tilde{\varphi})(\Phi) = \prod_{j=1}^{n(\Phi)} \left(\varepsilon^2(v(t_j^-) - v_j) \cdot \nu_j\right). \tag{4.5}$$

For $h > 0$ and $j = 0, \dots, n$ define $C_{h,j}(\Phi) = C_{h,j} \subset U \times \mathbb{R}^3$ to be the cube with side length h centred at $\tilde{\varphi}_j(\Phi)$. Further for $h > 0$ define,

$$C_h(\Phi) := \prod_{j=0}^{n} C_{h,j}(\Phi).$$

By the fact that the probability of finding a tree is less than the probability that initially there is a particles at the required initial position,

$$\begin{aligned} \hat{P}_t(\tilde{\varphi}^{-1}(C_h)) \leq \int_{C_{h,0}} f_0(x,v)\,\mathrm{d}x\,\mathrm{d}v \\ \times \prod_{j=1}^{n} N \int_{C_{h,j}} g_0(v)\,\mathrm{d}x\,\mathrm{d}v. \end{aligned} \tag{4.6}$$

Recalling that λ denotes the Lebesgue measure on $\mathcal{MT}$, by (4.5) it follows,

$$\lambda(\tilde{\varphi}^{-1}(C_h)) = \frac{h^{6(n+1)}}{\prod_{j=1}^{n}\left(\varepsilon^2(v(t_j^-) - v_j)\cdot \nu_j\right)}(1 + o(1)). \tag{4.7}$$

Hence combining (4.6) and (4.7) and recalling (2.1),

$$\begin{aligned} \frac{\hat{P}_t(\tilde{\varphi}^{-1}(C_h))}{\lambda(\tilde{\varphi}^{-1}(C_h))} \leq \frac{1}{h^6}\int_{C_{h,0}} f_0(x,v)\,\mathrm{d}x\,\mathrm{d}v \\ \times \prod_{j=1}^{n}\left(\frac{(v(t_j^-) - v_j)\cdot\nu_j}{h^6(1+o(1))}\int_{C_{h,j}} g_0(v)\,\mathrm{d}x\,\mathrm{d}v\right). \end{aligned} \tag{4.8}$$

Since $f_0 \in L^1(U \times \mathbb{R}^3)$ and $g_0 \in L^1(\mathbb{R}^3)$ let h tend to zero and the left hand side, which becomes $\hat{P}_t(\Phi)/\lambda(\Phi)$ in the limit, remains bounded. This completes the proof. □

We now prove the initial condition requirement on $\hat{P}_t$.

Lemma 4.10. *Under the assumptions and set up of theorem 4.6 we have*

$$\hat{P}_0(\Phi) = \zeta(\varepsilon) f_0(x_0, v_0)\mathbb{1}_{n(\Phi)=0}.$$

Proof. In the case $n(\Phi) > 0$, we have $\hat{P}_0(\Phi) = 0$. This is because the tree is free from initial overlap and the tree involves collisions happening at some positive time therefore the collisions cannot have occurred at time 0.
Now consider $n(\Phi) = 0$. In this situation the tree Φ contains only the root particle and the probability of finding the root at the given initial data (x_0, v_0) is given by $f_0(x_0, v_0)$. However this must be multiplied by a factor less than one because we must rule out situations that would give initial overlap of the root particle with a background particle. So we calculate the probability that there is no overlap. Firstly,

$$\begin{aligned} \mathbb{P}(|x_0 - x_1| > \varepsilon) = 1 - \mathbb{P}(|x_0 - x_1| < \varepsilon) = 1 - \int_{\mathbb{R}^3}\int_{|x_0 - x_1| < \varepsilon} g_0(\bar{v})\,\mathrm{d}x_1\,\mathrm{d}\bar{v} \\ = 1 - \frac{4}{3}\pi\varepsilon^3\int_{\mathbb{R}^3} g_0(\bar{v})\,\mathrm{d}\bar{v} = 1 - \frac{4}{3}\pi\varepsilon^3. \end{aligned}$$

Hence,

$$\mathbb{P}(|x_0 - x_j| > \varepsilon, \forall j = 1, \dots, N) = \mathbb{P}(|x_0 - x_1| > \varepsilon)^N = (1 - \frac{4}{3}\pi\varepsilon^3)^N = \zeta(\varepsilon),$$

as required. □

Before we prove the loss term lemma we require a few technical estimates to calculate the rate at which the root particle experiences a collision.

Definition 4.11. *Let $\Phi \in \mathcal{G}(\varepsilon)$. Define for $h > 0$,*

$$W_h(t) := \Big\{(\bar{x}, \bar{v}) \in U \times \mathbb{R}^3 : \exists (\nu', t') \in \mathbb{S}^2 \times (t, t+h) \\ \text{such that } x(t') + \varepsilon\nu' = \bar{x} + t'\bar{v} \text{ and } (v(t'^-) - \bar{v}) \cdot \nu' > 0\Big\}.$$

That is $W_h(t)$ is the set of initial points in $U \times \mathbb{R}^3$ for the background particles that lead to a collision with the root particle of Φ between the time t and $t+h$. Further define,

$$I_h(t) := \frac{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}. \tag{4.9}$$

From now on assume that the functions V and M in definition 4.5 satisfy, for any $0 < \varepsilon < 1$,

$$\varepsilon V(\varepsilon)^3 \le \frac{1}{8}, \tag{4.10}$$

and,

$$M(\varepsilon) \le \frac{1}{\sqrt{\varepsilon}}. \tag{4.11}$$

Lemma 4.12. *Recall definition 4.8. For ε sufficiently small, $\Phi \in \mathcal{G}(\varepsilon)$ and $t > \tau$,*

$$\lim_{h\to 0} \frac{1}{h} \hat{P}_t(\#(\omega \cap W_h(t)) \ge 2 \,|\, \Phi) = 0.$$

Proof. Note that by the inclusion exclusion principle the fact that the background particles are independent,

$$\begin{aligned} \hat{P}_t(\#(\omega \cap W_h(t)) \ge 2 \,|\, \Phi) &\le \sum_{1\le i<j\le N-n(\Phi)} \hat{P}_t((x_i, v_i) \in W_h(t) \text{ and } (x_j, v_j) \in W_h(t) \,|\, \Phi) \\ &\le N(N-1)\hat{P}_t((x_1, v_1) \in W_h(t) \text{ and } (x_2, v_2) \in W_h(t) \,|\, \Phi) \\ &= N(N-1)\hat{P}_t((x_1, v_1) \in W_h(t) \,|\, \Phi)^2. \end{aligned} \tag{4.12}$$

Recalling (4.9),

$$\hat{P}_t((x_1, v_1) \in W_h(t) \,|\, \Phi) = I_h(t). \tag{4.13}$$

Now we estimate the right hand side of (4.13) by estimating the numerator and denominator. Firstly by calculating the volume of the appropriate cylinder, for any $\bar{v} \in \mathbb{R}^3$,

$$\int_U \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \,\mathrm{d}\bar{x} \le \pi\varepsilon^2 \int_t^{t+h} |v(s) - \bar{v}| \,\mathrm{d}s. \tag{4.14}$$

Define

$$\beta := \int_{\mathbb{R}^3} g_0(v)(1 + |v|) \,\mathrm{d}v. \tag{4.15}$$

Note that by assumption (2.3), $\beta < \infty$. Since $\Phi \in \mathcal{G}(\varepsilon)$ it follows that $|v(t)| \le \mathcal{V}(\Phi) \le V(\varepsilon)$. Using these and (4.14) we estimate the numerator in (4.13),

$$\begin{aligned} \int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} &\le \int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} \\ &\le \int_{\mathbb{R}^3} g_0(\bar{v}) \pi\varepsilon^2 \int_t^{t+h} |v(s) - \bar{v}| \,\mathrm{d}s\,\mathrm{d}\bar{v} \le \pi\varepsilon^2 \int_{\mathbb{R}^3} g_0(\bar{v}) \int_t^{t+h} |v(s)| + |\bar{v}| \,\mathrm{d}s\,\mathrm{d}\bar{v} \\ &\le \pi\varepsilon^2 \int_{\mathbb{R}^3} g_0(\bar{v}) \int_t^{t+h} V(\varepsilon) + |\bar{v}| \,\mathrm{d}s\,\mathrm{d}\bar{v} \le \pi\varepsilon^2 \int_{\mathbb{R}^3} g_0(\bar{v}) h \left(V(\varepsilon) + |\bar{v}|\right) \,\mathrm{d}\bar{v} \\ &\le h\pi\varepsilon^2 \int_{\mathbb{R}^3} g_0(\bar{v}) \left(V(\varepsilon) + |\bar{v}|\right) \,\mathrm{d}\bar{v} \le h\pi\varepsilon^2 (V(\varepsilon) + \beta). \end{aligned} \tag{4.16}$$

Turning to the denominator of (4.9). Firstly note that,

$$\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} = \int_{U\times\mathbb{R}^3} g_0(\bar{v})\left(1-\mathbb{1}_{W_t(0)}(\bar{x},\bar{v})\right)\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}$$

$$= 1-\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_{W_t(0)}(\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}. \tag{4.17}$$

By using (4.14), $t\le T$ and the same estimates from the numerator estimate,

$$\int_{U\times R^3} g_0(\bar{v})\mathbb{1}_{W_t(0)}(\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}$$

$$\le \pi\varepsilon^2\int_{\mathbb{R}^3} g_0(\bar{v})\int_0^t |v(s)-\bar{v}|\,\mathrm{d}s\,\mathrm{d}\bar{v} \le \pi\varepsilon^2\int_{\mathbb{R}^3} g_0(\bar{v})\int_0^t \mathcal{V}(\Phi)+|\bar{v}|\,\mathrm{d}s\,\mathrm{d}\bar{v}$$

$$\le \pi\varepsilon^2\int_{\mathbb{R}^3} g_0(\bar{v})t\left(V(\varepsilon)+|\bar{v}|\right)\,\mathrm{d}\bar{v} \le \pi\varepsilon^2 T(V(\varepsilon)+\beta). \tag{4.18}$$

Hence for ε sufficiently small by (4.10),

$$\pi\varepsilon^2 T(V(\varepsilon)+\beta)\le 1/2,$$

so by (4.17) and (4.18),

$$\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} \ge 1/2. \tag{4.19}$$

Bounds for both the numerator and the denominator of (4.13) have been found in equations (4.16) and (4.19) respectively. Hence,

$$I_h(t)\le 2h\pi\varepsilon^2(V(\varepsilon)+\beta). \tag{4.20}$$

Substituting this into (4.12) and recalling (2.1),

$$\hat{P}_t(\#(\omega\cap W_h(t))\ge 2\,|\,\Phi) \le N(N-1)\times 4h^2\pi^2\varepsilon^4(V(\varepsilon)+\beta)^2$$
$$\le 4h^2\pi^2N^2\varepsilon^4(V(\varepsilon)+\beta)^2 \le 4h^2\pi^2(V(\varepsilon)+\beta)^2.$$

This gives finally that,

$$\lim_{h\to 0}\frac{1}{h}\left(\hat{P}_t(\#(\omega\cap W_h(t))\ge 2\,|\,\Phi)\right) \le \lim_{h\to 0}\frac{1}{h}\times 4h^2\pi^2(V(\varepsilon)+\beta)^2$$
$$\le 4\pi^2(V(\varepsilon)+\beta)^2\lim_{h\to 0}h = 0,$$

completing the proof of the lemma. □

The previous lemma shows that the rate of seeing two collisions in a short time converges to zero. We now show that the rate of seeing one collision converges to the required loss term. Before we do this we first estimate the error caused by re-collisions.

Definition 4.13. *For $\Phi\in\mathcal{G}(\varepsilon)$, $t>\tau$ and $h>0$ recall the definition of $W_h(t)$ in definition 4.11 and $\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})$ (4.1). Define, $B_{h,t}(\Phi)\subset U\times\mathbb{R}^3$,*

$$B_{h,t}(\Phi) := \left\{(\bar{x},\bar{v})\in U\times\mathbb{R}^3 : \mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})=0 \text{ and } \mathbb{1}_{W_h(t)}(\bar{x},\bar{v})=1\right\}.$$

Notice that $B_{h,t}(\Phi)$ is the set of all initial positions that a background particle can take such that it collides with the root once during $(0,t)$ and once during $(t,t+h)$.

Lemma 4.14. *For ε sufficiently small, $\Phi\in\mathcal{G}(\varepsilon)$, $t>\tau$ and $h>0$ sufficiently small there exists a $\hat{C}_2(\varepsilon)>0$ depending on t and Φ with $\hat{C}_2(\varepsilon)=o(1)$ as ε tends to zero such that,*

$$\int_{B_{h,t}(\Phi)} g_0(\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} = h\varepsilon^2\hat{C}_2(\varepsilon).$$

Proof. Recall that $(x(t), v(t))$ is the trajectory of the tagged particle defined by the tree Φ. $B_{h,t}(\Phi)$ is given by,

$$B_{h,t}(\Phi) = \big\{(\bar{x}, \bar{v}) : \exists s \in (0,t), \sigma \in (t, t+h), \nu_1, \nu_2 \in \mathbb{S}^2 \text{ such that } \bar{x} + s\bar{v} - x(s) = \varepsilon\nu_1,\ \bar{x} + \sigma\bar{v} - x(\sigma) = \varepsilon\nu_2 \text{ and } (v(s) - \bar{v}) \cdot \nu_1 > 0,\ (v(\sigma) - \bar{v}) \cdot \nu_2 > 0\big\}.$$

Define $\delta := \varepsilon^{1/3}$. We split the set $B_{h,t}(\Phi)$ into two parts, the first, denoted $B^{\delta}_{h,t}(\Phi)$, which considers $s \in (0, t-\delta]$ and the second, denoted $B^{2}_{h,t}(\Phi)$, which considers $s \in (t-\delta, t)$. We evaluate the bounds on these two sets separately.
Consider $s \in (0, t-\delta]$ and $\sigma \in (t, t+h)$ fixed. The conditions defined by $B_{h,t}$ require that,

$$\bar{v} = \frac{x(\sigma) - x(s)}{\sigma - s} + \frac{\varepsilon\nu_2 - \varepsilon\nu_1}{\sigma - s}.$$

For σ fixed this implies that $\bar{v}$ is contained in a cylinder of radius $2\varepsilon/\delta$ around the curve defined by $\frac{x(\sigma)-x(s)}{\sigma-s}$ for $s \in (0, t-\delta]$. Recalling the definition of $V(\varepsilon)$ from definition 4.5, taking $h \ll \varepsilon/V(\varepsilon)$ implies $h|v(\sigma)| \ll \varepsilon$ so the dependence on $\sigma \in (t, t+h)$ gives only a small perturbation around the curve defined by $\frac{x(t)-x(s)}{t-s}$ for $s \in (0, t-\delta]$. Hence $\bar{v}$ is contained in the cylinder with radius $4\varepsilon/\delta$ around the piecewise differentiable curve $r(s) := \frac{x(t)-x(s)}{t-s}$ for $s \in (0, t-\delta]$.
Denote this cylinder in $\mathbb{R}^3$ by $E = E(t, \Phi, \delta)$. We seek a bound on the volume of E, $|E|$. First consider the length of the curve r. For almost all $s \in (0, t-\delta)$,

$$\frac{\mathrm{d}}{\mathrm{d}s} r(s) = \frac{x(t) - x(s)}{(t-s)^2} + \frac{v(s)}{t-s}.$$

Hence,

$$|\frac{\mathrm{d}}{\mathrm{d}s} r(s)| \le \frac{|x(t) - x(s)|}{(t-s)^2} + \frac{|v(s)|}{t-s} \le \frac{3}{(t-s)^2} + \frac{V(\varepsilon)}{t-s}.$$

Thus the length of the curve is bounded by,

$$\int_0^{t-\delta} |\frac{\mathrm{d}}{\mathrm{d}s} r(s)| \,\mathrm{d}s \le \int_0^{t-\delta} \frac{3}{(t-s)^2} + \frac{V(\varepsilon)}{t-s} \,\mathrm{d}s = \frac{3}{\delta} - \frac{3}{t} - V(\varepsilon)(\log(\delta) - \log(t)).$$

Therefore for some $C > 0$,

$$|E| \le C \left(\frac{\varepsilon}{\delta}\right)^2 \left(\frac{3}{\delta} - \frac{3}{t} - V(\varepsilon)(\log(\delta) - \log(t))\right). \tag{4.21}$$

Noting that $x(\sigma) = x(t) + (\sigma - t)v(t)$, for $\bar{v}$ given, $(\bar{x}, \bar{v}) \in B_{h,t}(\Phi)$ requires that,

$$\begin{aligned} \bar{x} &= x(\sigma) - \sigma\bar{v} + \varepsilon\nu_2 = x(t) + (\sigma - t) - \sigma\bar{v} + \varepsilon\nu_2 \\ &= x(t) - t\bar{v} + (\sigma - t)(v(t) - \bar{v}) + \varepsilon\nu_2. \end{aligned}$$

Hence for $\bar{v}$ given $\bar{x}$ is contained in cylinder of radius ε and length $h|v(t) - \bar{v}|$. Denote this cylinder by $C(\bar{v})$. By (2.4) or (3.4), for constants C that change on each line,

$$\begin{aligned} \int_{B^{\delta}_{h,t}(\Phi)} g(\bar{v}) \,\mathrm{d}\bar{x} \,\mathrm{d}\bar{v} &\le \int_E g(\bar{v}) \int_{C(\bar{v})} \mathrm{d}\bar{x} \,\mathrm{d}\bar{v} \le C\varepsilon^2 h \int_E g(\bar{v}) |v(t) - \bar{v}| \,\mathrm{d}\bar{v} \\ &\le C\varepsilon^2 h \int_E g(\bar{v})(V(\varepsilon) + |\bar{v}|) \,\mathrm{d}\bar{v} \le C\varepsilon^2 h (V(\varepsilon) + 1)|E|. \end{aligned}$$

It remains to show that $(V(\varepsilon)+1)|E|$ is $o(1)$ as ε tends to zero. Recall (4.10), that $\delta = \varepsilon^{1/3}$ and (4.21),

$$\begin{aligned}(V(\varepsilon)+1)|E| &= (V(\varepsilon)+1) \times C\left(\frac{\varepsilon}{\delta}\right)^2 \left(\frac{3}{\delta} - \frac{3}{t} - V(\varepsilon)(\log(\delta) - \log(t))\right) \\ &\leq C\left(\frac{1}{2\varepsilon^{1/3}} + 1\right)\varepsilon^{4/3}\left(\frac{3}{\varepsilon^{1/3}} + \frac{1}{2\varepsilon^{1/3}}(|\log \varepsilon^{1/3}| + |\log t|)\right) \\ &\leq C\left(\frac{1}{2\varepsilon^{1/3}} + 1\right)\left(3\varepsilon + \frac{1}{2}\varepsilon\left(\frac{1}{3}|\log \varepsilon| + |\log t|\right)\right)\end{aligned}$$

as required.
Now consider the second part of $B_{h,t}(\Phi)$ for $s \in (t-\delta, t)$ denoted $B^2_{h,t}(\Phi)$. Since Φ is fixed, $t > \tau$ and $\delta = \varepsilon^{1/3}$ let ε sufficiently small such that $t - \delta > \tau$. Hence for $s \in (t-\delta, t)$, $v(s) = v(t)$. We change the velocity space coordinates so that $v(t) = 0$. If we require that a particle starting at $(\bar{x}, \bar{v})$ collides with the tagged particle in $(t-\delta, t)$ and again in $(t, t+h)$ we require in the new coordinates that either $\bar{v} = 0$ or that $|\bar{v}|$ is sufficiently large so that the background particle wraps round the torus having travelled at least distance $3/4$ (for ε sufficiently small) within time $(\delta + h)$ to re-collide with the tagged particle. That is,

$$|\bar{v}| \geq \frac{3}{4(\delta + h)}.$$

For $h \leq 1/4\delta$, this implies $|\bar{v}| \geq \frac{3}{5\delta}$. Changing back to the original coordinates, this means it is required that $\bar{v} = v(t)$ or $|\bar{v} - v(t)| \geq 3/5\delta$.
For a given $\bar{v}$, the same conditions as before on the $\bar{x}$ coordinate must hold and so $\bar{x}$ is in the cylinder $C(\bar{v})$. Recalling (2.3), (4.10), $\delta = \varepsilon^{1/3}$ and that $|v(t)| \leq V(\varepsilon) \leq 1/2\varepsilon^{-1/3} = 1/2\delta$ it follows for constants C that change on each line,

$$\begin{aligned}\int_{B^2_{h,t}(\Phi)} g_0(\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} &\leq \int_{\mathbb{R}^3\setminus B_{3/5\delta}(v(t))} g_0(\bar{v}) \int_{C(\bar{v})} \mathrm{d}\bar{x}\,\mathrm{d}\bar{v} \leq C\varepsilon^2 h \int_{\mathbb{R}^3\setminus B_{3/5\delta}(v(t))} g_0(\bar{v})|v(t) - \bar{v}|\,\mathrm{d}\bar{v} \\ &\leq C\varepsilon^2 h \int_{\mathbb{R}^3\setminus B_{3/5\delta}(v(t))} g_0(\bar{v})(V(\varepsilon) + |\bar{v}|)\,\mathrm{d}\bar{v} \leq C\varepsilon^2 h \int_{\mathbb{R}^3\setminus B_{1/10\delta}(0)} g_0(\bar{v})(V(\varepsilon) + |\bar{v}|)\,\mathrm{d}\bar{v} \\ &\leq C\varepsilon^2 h \int_{\mathbb{R}^3\setminus B_{1/10\delta}(0)} g_0(\bar{v})(100\delta^2|\bar{v}|^2 V(\varepsilon) + 10\delta|\bar{v}|^2)\,\mathrm{d}\bar{v} \leq C\varepsilon^2 h(10\delta^2 V(\varepsilon) + \delta) \\ &\leq C\varepsilon^2 h\left(5\frac{\varepsilon^{2/3}}{\varepsilon^{1/3}} + \varepsilon^{1/3}\right) = C\varepsilon^2 h \times \varepsilon^{1/3},\end{aligned}$$

as required. Since $B_{h,t}(\Phi) = B^\delta_{h,t}(\Phi) \cup B^2_{h,t}(\Phi)$ the proof of the lemma is complete.

□

Lemma 4.15. *For ε sufficiently small, $\Phi \in \mathcal{G}(\varepsilon)$, $t > \tau$ and $\hat{C}_2(\varepsilon)$ as in the above lemma,*

$$\begin{aligned}&\lim_{h\to 0} \frac{1}{h}\hat{P}_t(\#(\omega \cap W_h(t)) = 1 \,|\, \Phi) \\ &\qquad = (1 - \gamma(t))\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(t) - \bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu - \hat{C}_2(\varepsilon)}{\int_{U\times R^3} g_0(\bar{v})\mathbb{1}^\varepsilon_t[\Phi](\bar{x}, \bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}.\end{aligned}$$

Proof. Since the initial data for each background particle is independent of the other background particles,

$$\begin{aligned}
\hat{P}_t(\#(\omega\cap W_h(t)) = 1 \,|\, \Phi) &= \sum_{i=1}^{N-n(\Phi)} \hat{P}_t\Big((x_i, v_i) \in W_h(t) \text{ and } (x_1, v_1), \dots, \\
&\qquad (x_{i-1}, v_{i-1}), (x_{i+1}, v_{i+1}), \dots, (x_{N-n(\Phi)}, v_{N-n(\Phi)}) \notin W_h(t) \,|\, \Phi\Big) \\
&= (N - n(\Phi))\hat{P}_t((x_1, v_1) \in W_h(t) \,|\, \Phi)\hat{P}_t((x_2, v_2) \notin W_h(t) \,|\, \Phi)^{N-n(\Phi)-1} \\
&= (N - n(\Phi))I_h(t)\,(1 - I_h(t))^{N-n(\Phi)-1} \\
&= (N - n(\Phi))I_h(t) \sum_{j=0}^{N-n(\Phi)-1} (-1)^j \binom{N - n(\Phi) - 1}{j} I_h(t)^j \\
&= (N - n(\Phi)) \sum_{j=0}^{N-n(\Phi)-1} (-1)^j \binom{N - n(\Phi) - 1}{j} I_h(t)^{j+1}. \qquad (4.22)
\end{aligned}$$

By (4.20),

$$\lim_{h\to 0} \frac{1}{h} I_h(t)^2 = 0.$$

Hence dividing (4.22) by h and taking h to zero we see that all terms in the sum for $j \geq 1$ tend to zero, leaving only the contribution from the term $j = 0$. Hence,

$$\begin{aligned}
&\lim_{h\to 0} \frac{1}{h} \hat{P}_t(\#(\omega \cap W_h(t)) = 1 \,|\, \Phi) \\
&= \lim_{h\to 0} \frac{1}{h}(N - n(\Phi)) \sum_{j=0}^{N-n(\Phi)-1} (-1)^j \binom{N - n(\Phi) - 1}{j} I_h(t)^{j+1} = \lim_{h\to 0} \frac{1}{h}(N - n(\Phi))I_h(t). \qquad (4.23)
\end{aligned}$$

It remains to investigate,

$$\lim_{h\to 0} \frac{1}{h} I_h(t) = \lim_{h\to 0} \frac{1}{h} \frac{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}.$$

For $B_{h,t}(\Phi)$ as defined in definition 4.13 we have,

$$\begin{aligned}
&\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} \\
&\qquad = \int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \,\mathrm{d}\bar{v} - \int_{B_{h,t}(\Phi)} g_0(\bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}
\end{aligned}$$

It then follows from lemma 4.14,

$$\begin{aligned}
\lim_{h\to 0} \frac{1}{h} I_h(t) &= \lim_{h\to 0} \frac{1}{h} \frac{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_{W_h(t)}(\bar{x}, \bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}} \\
&= \lim_{h\to 0} \frac{1}{h} \left(\frac{h\varepsilon^2 \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v}) [(v(t) - \bar{v}) \cdot \nu]_+ \,\mathrm{d}\bar{v}\,\mathrm{d}\nu - h\varepsilon^2 \hat{C}_2(\varepsilon)}{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}} \right) \\
&= \varepsilon^2 \frac{\int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v}) [(v(t) - \bar{v}) \cdot \nu]_+ \,\mathrm{d}\bar{v}\,\mathrm{d}\nu - \hat{C}_2(\varepsilon)}{\int_{U\times\mathbb{R}^3} g_0(\bar{v}) \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}) \,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}.
\end{aligned}$$

Substituting this into (4.23),

$$\begin{aligned}\lim_{h\to 0}\frac{1}{h}\hat{P}_t(\#(\omega\cap W_h(t))=1\,|\,\Phi)&=\lim_{h\to 0}\frac{1}{h}(N-n(\Phi))I_h(t)\\&=(N-n(\Phi))\varepsilon^2\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3}g_0(\bar{v})[(v(t)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu-\hat{C}_2(\varepsilon)}{\int_{U\times\mathbb{R}^3}g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}\\&=(1-\gamma(t))\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3}g_0(\bar{v})[(v(t)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu-\hat{C}_2(\varepsilon)}{\int_{U\times\mathbb{R}^3}g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}},\end{aligned}$$

which proves the lemma. □

With these estimates we can now prove that the loss term of (4.2) holds.

Lemma 4.16. *Under the assumptions and set up of theorem 4.6, for $t>\tau$,*

$$\partial_t\hat{P}_t(\Phi)=(1-\gamma(t))\hat{\mathcal{Q}}_t^-[\hat{P}_t](\Phi).$$

Proof. We calculate, for $t>\tau$,

$$\partial_t\hat{P}_t(\Phi)=\lim_{h\to 0}\frac{1}{h}\left(\hat{P}_{t+h}(\Phi)-\hat{P}_t(\Phi)\right).$$

Noting that,

$$\hat{P}_{t+h}(\Phi)=\left(1-\hat{P}_t\left(\#(\omega\cap W_h(t))>0\,|\,\Phi\right)\right)\hat{P}_t(\Phi).$$

It follows,

$$\frac{1}{h}\left(\hat{P}_{t+h}(\Phi)-\hat{P}_t(\Phi)\right)=-\frac{1}{h}\hat{P}_t(\#(\omega\cap W_h(t))>0\,|\,\Phi)\hat{P}_t(\Phi).$$

Using lemma 4.12 and lemma 4.15,

$$\begin{aligned}\partial_t\hat{P}_t(\Phi)&=\lim_{h\to 0}\frac{1}{h}\left(\hat{P}_{t+h}(\Phi)-\hat{P}_t(\Phi)\right)\\&=-\lim_{h\to 0}\frac{1}{h}\hat{P}_t(\#(\omega\cap W_h(t))>0\,|\,\Phi)\hat{P}_t(\Phi)=-\hat{P}_t(\Phi)\lim_{h\to 0}\frac{1}{h}\hat{P}_t(\#(\omega\cap W_h(t))=1\,|\,\Phi)\\&=-\hat{P}_t(\Phi)(1-\gamma(t))\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3}g_0(\bar{v})[(v(t)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu-\hat{C}_2(\varepsilon)}{\int_{U\times R^3}g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}=(1-\gamma(t))\hat{\mathcal{Q}}_t^-[\hat{P}_t](\Phi),\end{aligned}$$

which proves the lemma. □

We next move to proving the gain term in (4.2).

Lemma 4.17. *Under the assumptions and set up of theorem 4.6, for $n(\Phi)\geq 1$*

$$\hat{P}_\tau(\Phi)=(1-\gamma(\tau))\hat{P}_\tau(\bar{\Phi})\frac{g_0(v')[(v(\tau^-)-v')\cdot\nu]_+}{1-\pi\varepsilon^2\int_0^\tau|v(s)-v'|\,\mathrm{d}s+\hat{C}_1\varepsilon^3}.$$

For some $\hat{C}_1>0$ depending only on Φ.

Proof. Firstly,

$$\hat{P}_\tau(\Phi)=\hat{P}_\tau(\Phi\cap\bar{\Phi})=\hat{P}_\tau(\Phi\,|\,\bar{\Phi})\hat{P}_\tau(\bar{\Phi}).$$

It remains to show that

$$\hat{P}_\tau(\Phi\,|\,\bar{\Phi})=(1-\gamma(\tau))\frac{g_0(v')[(v(\tau^-)-v')\cdot\nu]_+}{1-\pi\varepsilon^2\int_0^\tau|v(s)-v'|\,\mathrm{d}s+\hat{C}_1\varepsilon^3}. \tag{4.24}$$

We do this by proving upper and lower bounds. For $h\geq 0$ define,

$$U_h:=\{\Psi\in\mathcal{MT}:\bar{\Psi}_{\tau(\Phi)}=\bar{\Phi}_{\tau(\Phi)}\text{ and }\Psi\in B_h(\Phi)\}.$$

Note that $U_0:=\{\Phi\}$. Then by lemma 4.9,

$$\hat{P}_\tau(\Phi\,|\,\bar{\Phi})=\lim_{h\to 0}h^{-6}\hat{P}_\tau(U_h\,|\,\bar{\Phi}). \tag{4.25}$$

For $\Psi \in U_h$ define $V_h(\Psi) \in U \times \mathbb{R}^3$ to be the initial position of the background particle that leads to the final collision of Ψ and define $V_h \subset U \times \mathbb{R}^3$ by,

$$V_h = \cup_{\Psi \in U_h} V_h(\Psi).$$

Note that $V_0 = \{(x(\tau) + \varepsilon\nu - \tau v', v')\}$ that is V_0 contains only the initial point of the background particle that gives the final collision in Φ. Then by a change of coordinates, recalling (4.4),

$$\begin{aligned}
\hat{P}_\tau(U_h \,|\, \bar{\Phi}) &\leq \sum_{i=1}^{N-(n(\Phi)-1)} \hat{P}_\tau((x_i, v_i) \in V_h \,|\, \bar{\Phi}_\tau)\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+ \\
&= (N - (n(\Phi) - 1))\varepsilon^2 \hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi})[(v(\tau^-) - v') \cdot \nu]_+ \\
&= (1 - \gamma(\tau))\hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi})[(v(\tau^-) - v') \cdot \nu]_+. \qquad (4.26)
\end{aligned}$$

By absolute continuity of $\hat{P}_\tau$,

$$\lim_{h \to 0} h^{-6} \hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi}) = \hat{P}_\tau((x_1, v_1) \in V_0 \,|\, \bar{\Phi}).$$

Combining these into (4.25),

$$\begin{aligned}
\hat{P}_\tau(\Phi \,|\, \bar{\Phi}) &= \lim_{h \to 0} h^{-6} \hat{P}_\tau(U_h \,|\, \bar{\Phi}) \\
&\leq \lim_{h \to 0} h^{-6}(1 - \gamma(\tau))\hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi})[(v(\tau^-) - v') \cdot \nu]_+ \\
&= (1 - \gamma(\tau))\hat{P}_\tau((x_1, v_1) \in V_0 \,|\, \bar{\Phi}_\tau)[(v(\tau^-) - v') \cdot \nu]_+. \qquad (4.27)
\end{aligned}$$

Next consider the lower bound. By the inclusion-exclusion principle,

$$\begin{aligned}
\hat{P}_\tau(U_h \,|\, \bar{\Phi}) \geq &\sum_{i=1}^{N-(n(\Phi)-1)} \hat{P}_\tau((x_i, v_i) \in V_h \,|\, \bar{\Phi})\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+ \\
&- \sum_{1 \leq i < j \leq N-(n(\Phi)-1)} \hat{P}_\tau((x_i, v_i), (x_j, v_j) \in V_h \,|\, \bar{\Phi})\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+. \qquad (4.28)
\end{aligned}$$

As in (4.26) it follows,

$$\begin{aligned}
&\sum_{i=1}^{N-(n(\Phi)-1)} \hat{P}_\tau((x_i, v_i) \in V_h \,|\, \bar{\Phi})\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+ \\
&\qquad = (1 - \gamma(\tau))\hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi})[(v(\tau^-) - v') \cdot \nu]_+.
\end{aligned}$$

Further,

$$\begin{aligned}
&\sum_{1 \leq i < j \leq N-(n(\Phi)-1)} \hat{P}_\tau((x_i, v_i), (x_j, v_j) \in V_h \,|\, \bar{\Phi})\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+ \\
&\qquad \leq N(N-1)\hat{P}_\tau((x_1, v_1), (x_2, v_2) \in V_h \,|\, \bar{\Phi})\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+ \\
&\qquad = (N-1)\hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi})^2[(v(\tau^-) - v') \cdot \nu]_+.
\end{aligned}$$

By the absolute continuity of $\hat{P}_t$ this implies,

$$\lim_{h \to 0} \frac{1}{h^6} \sum_{1 \leq i < j \leq N-(n(\Phi)-1)} \hat{P}_\tau((x_i, v_i), (x_j, v_j) \in V_h \,|\, \bar{\Phi})\varepsilon^2[(v(\tau^-) - v') \cdot \nu]_+ = 0.$$

Hence by (4.28),

$$\begin{aligned}
\hat{P}_\tau(\Phi \,|\, \bar{\Phi}) &= \lim_{h \to 0} h^{-6} \hat{P}_\tau(U_h \,|\, \bar{\Phi}) \\
&\geq \lim_{h \to 0} h^{-6}(1 - \gamma(\tau))\hat{P}_\tau((x_1, v_1) \in V_h \,|\, \bar{\Phi})[(v(\tau^-) - v') \cdot \nu]_+ \\
&= (1 - \gamma(\tau))\hat{P}_\tau((x_1, v_1) \in V_0 \,|\, \bar{\Phi})[(v(\tau^-) - v') \cdot \nu]_+. \qquad (4.29)
\end{aligned}$$

Recalling that we need to prove (4.24) to prove the lemma, we see that with (4.27) and (4.29) we now need only to show that,

$$\hat{P}_\tau((x_1, v_1) \in V_0 \,|\, \bar{\Phi}) = \frac{g_0(v')}{1 - \pi\varepsilon^2 \int_0^\tau |v(s) - v'| \,\mathrm{d}s + \hat{C}_1\varepsilon^3}.$$

That is,

$$\hat{P}_\tau\left((x_1, v_1) = (x(\tau) + \varepsilon\nu - \tau v', v') \,|\, \bar{\Phi}\right) = \frac{g_0(v')}{1 - \pi\varepsilon^2 \int_0^\tau |v(s) - v'| \,\mathrm{d}s + \hat{C}_1\varepsilon^3}. \tag{4.30}$$

Since we are conditioning on $\bar{\Phi}$ occurring there is a region of U that we must rule out for the initial position of the background particle if we know that is has velocity v'. That is to say there is a region of U where we know the background particle cannot have started with initial velocity v' because if it did it would have interfered with $\bar{\Phi}$. Denote this region of U by Δ. To calculate the volume of Δ we imagine it as cylinders to get, for fixed incoming velocity v',

$$|\Delta| = \pi\varepsilon^2 \int_0^\tau |v(s) - v'| \,\mathrm{d}s - \hat{C}_1\varepsilon^3, \tag{4.31}$$

where the first term calculates the volume of each cylinder as the root particle changes direction and the second term subtracts the over-estimate error caused counting certain parts twice as the particle changes direction.

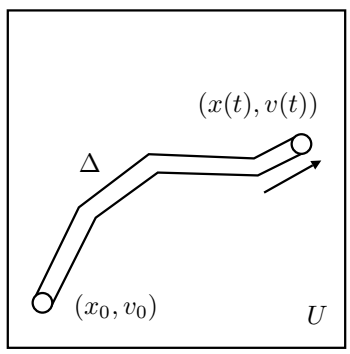


FIGURE 3. In the case $v' = 0$ we are calculating the volume of Δ, since we know the background particle cannot start in Δ. For $v' \neq 0$ the cylinders get shifted but the principle is the same. (Diagram not to scale)

Therefore, recalling that $|U| = 1$,

$$|U \setminus \Delta| = 1 - \left(\pi\varepsilon^2 \int_0^\tau |v(s) - v'| \,\mathrm{d}s - \hat{C}_1\varepsilon^3\right),$$

and this together with the fact that the velocity of the background particle has initial distribution g_0 gives the required (4.30). □

Proof of theorem 4.6. Combining lemmas 4.10, 4.16 and 4.17 gives the required proof of the theorem. □

5. Convergence

Having proven the existence of the idealized distribution P_t and shown that the empirical distribution $\hat{P}_t$ solves the appropriate equation, we seek to show the convergence results that will help prove our main result. Following [26], the idea is to establish a differential inequality in (5.3). In combination with the fact that P_t is a probability measure and that $\lim_{\varepsilon\to 0} P_t(\mathcal{G}(\varepsilon)) = 1$ in proposition 5.5 the inequality delivers the convergence result theorem 5.8. The main theorem 2.3

is a direct consequence. We first introduce some notation. Recall (4.1) and (4.3). For $\varepsilon > 0$, $\Phi \in \mathcal{G}(\varepsilon)$, $t \in [0,T]$, define the following,

$$
\begin{aligned}
\eta_t^\varepsilon(\Phi) &:= \int_{U\times\mathbb{R}^3} g_0(\bar{v})(1 - \mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v},\\
R_t^\varepsilon(\Phi) &:= \zeta(\varepsilon)P_t(\Phi),\\
L(\Phi) &:= -\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu,\\
C(\Phi) &:= 2\sup_{t\in[0,T]}\left\{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(t)-\bar{v})\cdot\nu]_+\right\}\\
\rho_t^{\varepsilon,0}(\Phi) &:= \eta_t^\varepsilon(\Phi)C(\Phi)t.
\end{aligned}
\tag{5.1}
$$

Further for $k \geq 1$ define,

$$
\rho_t^{\varepsilon,k}(\Phi) := (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Phi) + \rho_t^{\varepsilon,0}(\Phi) + \varepsilon.
$$

Note that this implies that for $k \geq 1$,

$$
\rho_t^{\varepsilon,k}(\Phi) = (1-\varepsilon)^k\rho_t^{\varepsilon,0}(\Phi) + (\rho_t^{\varepsilon,0}(\Phi)+\varepsilon)\sum_{j=1}^{k}(1-\varepsilon)^{k-j}. \tag{5.2}
$$

Finally define,

$$
\hat{\rho}_t^\varepsilon(\Phi) := \rho_t^{\varepsilon,n(\Phi)}(\Phi).
$$

Proposition 5.1. *For ε sufficiently small, $\Phi \in \mathcal{G}(\varepsilon)$ and $t \in [0,T]$,*

$$
\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi) \geq -\hat{\rho}_t^\varepsilon(\Phi)R_t^\varepsilon(\Phi). \tag{5.3}
$$

To prove this proposition we use a number of lemmas.

Lemma 5.2. *For $\Phi \in \mathcal{G}(\varepsilon)$ and $t \geq \tau$,*

$$
\begin{aligned}
\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi) \geq \exp\left(L(\Phi)\int_\tau^t (1+2\eta_s^\varepsilon(\Phi))\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Phi) - R_\tau^\varepsilon(\Phi))\\
+ 2\eta_t^\varepsilon(\Phi)L(\Phi)R_t^\varepsilon(\Phi)\int_\tau^t \exp\left(2\eta_s^\varepsilon(\Phi)(t-s)L(\Phi)\right)\,\mathrm{d}s.
\end{aligned}
$$

Proof. For $t = \tau$ the result holds trivially. For $t > \tau$, by theorem 3.1 and theorem 4.6,

$$
\partial_t\left(\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi)\right) = (1-\gamma(t))\hat{L}_t(\Phi)\hat{P}_t(\Phi) - L(\Phi)R_t^\varepsilon(\Phi), \tag{5.4}
$$

where

$$
\hat{L}_t(\Phi) := -\left(\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu - \hat{C}_2(\varepsilon)}{\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}\right).
$$

Recalling (4.31) and (4.15), for ε sufficiently small,

$$
\begin{aligned}
\int_{U\times\mathbb{R}^3} g_0(\bar{v})(1-\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} &\leq \int_{\mathbb{R}^3} g_0(\bar{v})\left(\pi\varepsilon^2\int_0^t |v(s)-\bar{v}|\,\mathrm{d}s - \hat{C}_1\varepsilon^3\right)\mathrm{d}\bar{v}\\
&\leq \int_{\mathbb{R}^3} g_0(\bar{v})\left(\pi\varepsilon^2 T(\mathcal{V}(\Phi)+|\bar{v}|) - \hat{C}_1\varepsilon^3\right)\mathrm{d}\bar{v} \leq \int_{\mathbb{R}^3} g_0(\bar{v})\left(\pi\varepsilon^2 T(V(\varepsilon)+|\bar{v}|)\right)\mathrm{d}\bar{v}\\
&\leq \pi\varepsilon^2 T\int_{\mathbb{R}^3} g_0(\bar{v})\left(V(\varepsilon)+|\bar{v}|\right)\mathrm{d}\bar{v} \leq \pi\varepsilon^2 T(V(\varepsilon)+\beta) < 1/2.
\end{aligned}
\tag{5.5}
$$

Noting that for $0 \leq z \leq 1/2$, $\sum_{i=0}^\infty z^i \leq 2$ which gives,

$$
\frac{1}{1-z} = \sum_{i=0}^\infty z^i = 1 + z\left(\sum_{i=0}^\infty z^i\right) \leq 1 + 2z.
$$

It follows,

$$\frac{1}{\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}} = \frac{1}{1-\int_{U\times\mathbb{R}^3} g_0(\bar{v})(1-\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}$$

$$\leq 1 + 2\left(\int_{U\times\mathbb{R}^3} g_0(\bar{v})(1-\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}\right) = 1 + 2\eta_t^\varepsilon(\Phi). \tag{5.6}$$

This gives that,

$$\begin{aligned}(1-\gamma(t))&\left(\frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu - \hat{C}_2(\varepsilon)}{\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}\right)\\ &\leq \frac{\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu}{\int_{U\times\mathbb{R}^3} g_0(\bar{v})\mathbb{1}_t^\varepsilon[\Phi](\bar{x},\bar{v})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}}\\ &\leq (1+2\eta_t^\varepsilon(\Phi))\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(\tau)-\bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu.\end{aligned}$$

Finally giving that,

$$(1-\gamma(t))\hat{L}_t(\Phi) \geq (1+2\eta_t^\varepsilon(\Phi))L(\Phi).$$

Substituting this into (5.4),

$$\begin{aligned}\partial_t\left(\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi)\right) &= (1-\gamma(t))\hat{L}_t(\Phi)\hat{P}_t(\Phi) - L(\Phi)R_t^\varepsilon(\Phi)\\ &\geq (1+2\eta_t^\varepsilon(\Phi))L(\Phi)\hat{P}_t(\Phi) - L(\Phi)R_t^\varepsilon(\Phi)\\ &= (1+2\eta_t^\varepsilon(\Phi))L(\Phi)\left(\hat{P}_t(\Phi) - R_t^\varepsilon(\Phi)\right) + 2\eta_t^\varepsilon(\Phi)L(\Phi)R_t^\varepsilon(\Phi).\end{aligned} \tag{5.7}$$

For fixed Φ this is simply a $1d$ ODE in t. If $y:[\tau,\infty)\to\mathbb{R}$ satisfies,

$$\begin{cases}\frac{d}{dt}y(t) \geq a(t)y(t) + b(t),\\ y(\tau) = y_0.\end{cases}$$

Then it follows by the variation of constants,

$$y(t) \geq \exp\left(\int_\tau^t a(s)\,\mathrm{d}s\right)y_0 + \int_\tau^t \exp\left(\int_s^t a(\sigma)\,\mathrm{d}\sigma\right)b(s)\,\mathrm{d}s.$$

Applying this to (5.7),

$$\begin{aligned}\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi) \geq \exp&\left(\int_\tau^t (1+2\eta_s^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Phi) - R_\tau^\varepsilon(\Phi))\\ &+ \int_\tau^t \exp\left(\int_s^t (1+2\eta_\sigma^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}\sigma\right)2\eta_s^\varepsilon(\Phi)L(\Phi)R_s^\varepsilon(\Phi)\,\mathrm{d}s.\end{aligned} \tag{5.8}$$

Recall the definition of $\mathbb{1}_t^\varepsilon[\Phi]$ (4.1) and note that it is non-increasing. Hence $\eta_t^\varepsilon(\Phi)$ is non-decreasing. So for $\tau \leq \sigma \leq t$, $\eta_\sigma^\varepsilon(\Phi) \leq \eta_t^\varepsilon(\Phi)$. Recalling that $L(\Phi)$ is non-positive, (5.8) becomes,

$$\begin{aligned}\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi) &\geq \exp\left(\int_\tau^t (1+2\eta_s^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Phi) - R_\tau^\varepsilon(\Phi))\\ &\quad + 2\eta_t^\varepsilon(\Phi)\int_\tau^t \exp\left(\int_s^t (1+2\eta_\sigma^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}\sigma\right)L(\Phi)R_s^\varepsilon(\Phi)\,\mathrm{d}s\\ &\geq \exp\left(\int_\tau^t (1+2\eta_s^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Phi) - R_\tau^\varepsilon(\Phi))\\ &\quad + 2\eta_t^\varepsilon(\Phi)L(\Phi)\int_\tau^t \exp\left((1+2\eta_s^\varepsilon(\Phi))(t-s)\,L(\Phi)\right)R_s^\varepsilon(\Phi)\,\mathrm{d}s.\end{aligned} \tag{5.9}$$

Finally for $t > \tau$,

$$\partial_t R_t^\varepsilon(\Phi) = \zeta(\varepsilon)\partial_t P_t(\Phi) = \zeta(\varepsilon)P_t(\Phi)L(\Phi) = R_t^\varepsilon(\Phi)L(\Phi).$$

Hence for $\tau \le s \le t$,

$$R_t^\varepsilon(\Phi) = \exp\left((t-s)L(\Phi)\right) R_s^\varepsilon(\Phi).$$

Which implies,

$$R_s^\varepsilon(\Phi) = \exp\left(-(t-s)L(\Phi)\right) R_t^\varepsilon(\Phi). \tag{5.10}$$

Substituting this into (5.9),

$$\begin{aligned}
\hat{P}_t^\varepsilon(\Phi) - R_t^\varepsilon(\Phi) &\ge \exp\left(\int_\tau^t (1+2\eta_s^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Phi) - R_\tau^\varepsilon(\Phi)) \\
&\quad + 2\eta_t^\varepsilon(\Phi)L(\Phi)\int_\tau^t \exp\Big((1+2\eta_s^\varepsilon(\Phi))(t-s)L(\Phi)\Big) \\
&\quad \times \exp\left(-(t-s)L(\Phi)\right) R_t^\varepsilon(\Phi)\,\mathrm{d}s \\
&\ge \exp\left(\int_\tau^t (1+2\eta_s^\varepsilon(\Phi))L(\Phi)\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Phi) - R_\tau^\varepsilon(\Phi)) \\
&\quad + 2\eta_t^\varepsilon(\Phi)L(\Phi)R_t^\varepsilon(\Phi)\int_\tau^t \exp\left(2\eta_s^\varepsilon(\Phi)(t-s)L(\Phi)\right)\,\mathrm{d}s.
\end{aligned}$$

This completes the proof of the lemma. □

Lemma 5.3. *For $\Phi \in \mathcal{G}(\varepsilon)$ and $t \ge \tau$,*

$$2\eta_t^\varepsilon(\Phi)L(\Phi)\int_\tau^t \exp\left(2\eta_s^\varepsilon(\Phi)(t-s)L(\Phi)\right)\,\mathrm{d}s \ge -\rho_t^{\varepsilon,0}(\Phi).$$

Proof. Since $L(\Phi) \le 0$,

$$\int_\tau^t \exp\left(2\eta_s^\varepsilon(\Phi)(t-s)L(\Phi)\right)\,\mathrm{d}s \le t - \tau \le t.$$

Recalling (5.1),

$$-L(\Phi) \le \frac{C(\Phi)}{2}.$$

Hence combining these,

$$2L(\Phi)\int_\tau^t \exp\left(2\eta_s^\varepsilon(\Phi)(t-s)L(\Phi)\right)\,\mathrm{d}s \ge -C(\Phi)t.$$

Multiplying both sides by $\eta_t^\varepsilon(\Phi)$ gives the required identity. □

Lemma 5.4. *For ε sufficiently small, for any $\Phi \in \mathcal{G}(\varepsilon)$ and any $t \in [0,T]$,*

$$1 - \frac{1-\gamma(t)}{1-|\Delta|} \le \varepsilon.$$

Proof. We show that for ε sufficiently small,

$$\frac{1}{\varepsilon}\left(1 - \frac{1-\gamma(t)}{1-|\Delta|}\right) \le 1.$$

Note,

$$|\Delta| = \pi\varepsilon^2\int_0^\tau |v(s) - v'|\,\mathrm{d}s - \hat{C}\varepsilon^3 \le \pi\varepsilon^2\int_0^\tau |v(s) - v'|\,\mathrm{d}s \le \pi\varepsilon^2 2TV(\varepsilon).$$

Recalling (4.11),

$$1-\frac{1-\gamma(t)}{1-|\Delta|}=\frac{1-|\Delta|}{1-|\Delta|}-\frac{1-\gamma(t)}{1-|\Delta|}=\frac{\gamma(t)-|\Delta|}{1-|\Delta|}$$
$$\leq\frac{\gamma(t)}{1-|\Delta|}\leq\frac{\varepsilon^2 n(\Phi)}{1-2\varepsilon^2\pi TV(\varepsilon)}\leq\frac{\varepsilon^2 M(\varepsilon)}{1-2\varepsilon^2\pi TV(\varepsilon)}\leq\frac{\varepsilon^{3/2}}{1-2\varepsilon^2\pi TV(\varepsilon)}.$$

Therefore

$$\frac{1}{\varepsilon}\left(1-\frac{1-\gamma(t)}{1-|\Delta|}\right)\leq\frac{\varepsilon^{1/2}}{1-2\varepsilon^2\pi TV(\varepsilon)},$$

which, by recalling (4.10), converges to zero as ε converges to zero. Hence for ε sufficiently small the right hand side is less than 1.

□

Proof of proposition 5.1. We prove by induction on the degree of Φ. Firstly we show that the proposition holds for $\Phi\in\mathcal{T}_0\cap\mathcal{G}(\varepsilon)$. Now if $\Phi\in\mathcal{T}_0\cap\mathcal{G}(\varepsilon)$ it follows that $\tau=0$ and hence,

$$\hat{P}^\varepsilon_\tau(\Phi)=\hat{P}^\varepsilon_0(\Phi)=\zeta(\varepsilon)f_0(x_0,v_0)=\zeta(\varepsilon)P_0(\Phi)=R^\varepsilon_\tau(\Phi).$$

By lemma 5.2 and 5.3 for $t\geq 0$,

$$\hat{P}^\varepsilon_t(\Phi)-R^\varepsilon_t(\Phi)\geq 2\eta^\varepsilon_t(\Phi)L(\Phi)R^\varepsilon_t(\Phi)\int_\tau^t\exp\left(2\eta^\varepsilon_s(\Phi)(t-s)L(\Phi)\right)\,\mathrm{d}s\geq-\rho^{\varepsilon,0}_t(\Phi)R^\varepsilon_t(\Phi)$$
$$=-\hat{\rho}^\varepsilon_t(\Phi)R^\varepsilon_t(\Phi).$$

Proving the proposition in the base case. Now suppose that the proposition holds true for all trees $\Phi\in\mathcal{T}_{k-1}\cap\mathcal{G}(\varepsilon)$ for some $k\geq 1$ and let $\Psi\in\mathcal{T}_k\cap\mathcal{G}(\varepsilon)$. For $t<\tau$ the proposition holds trivially, so we consider $t\geq\tau$. Now recall that,

$$\hat{P}^\varepsilon_\tau(\Psi)=\frac{1-\gamma(t)}{1-|\Delta|}\hat{P}^\varepsilon_\tau(\bar{\Psi})g_0(v')[(v(\tau^-)-v')\cdot\nu]_+,$$

and,

$$R^\varepsilon_\tau(\Psi)=R^\varepsilon_\tau(\bar{\Psi})g_0(v')[(v(\tau^-)-v')\cdot\nu]_+.$$

Further since $\bar{\Psi}\in\mathcal{T}_{k-1}$ we know by our inductive assumption that the proposition holds for $\bar{\Psi}$, which implies.

$$\hat{P}^\varepsilon_t(\bar{\Psi})\geq R^\varepsilon_t(\bar{\Psi})-\hat{\rho}^\varepsilon_t(\bar{\Psi})R^\varepsilon_t(\bar{\Psi}).$$

Hence by the estimate in lemma 5.4 for ε sufficiently small,

$$\hat{P}^\varepsilon_\tau(\Psi)-R^\varepsilon_\tau(\Psi)=g_0(v')[(v(\tau^-)-v')\cdot\nu]_+\left(\frac{1-\gamma(t)}{1-|\Delta|}\hat{P}^\varepsilon_\tau(\bar{\Psi})-R^\varepsilon_\tau(\bar{\Psi})\right)$$
$$\geq g_0(v')[(v(\tau^-)-v')\cdot\nu]_+\left((1-\varepsilon)\hat{P}^\varepsilon_\tau(\bar{\Psi})-R^\varepsilon_\tau(\bar{\Psi})\right)$$
$$\geq g_0(v')[(v(\tau^-)-v')\cdot\nu]_+\left((1-\varepsilon)(R^\varepsilon_\tau(\bar{\Psi})-\hat{\rho}_t(\bar{\Psi})R^\varepsilon_\tau(\bar{\Psi}))-R^\varepsilon_\tau(\bar{\Psi})\right)$$
$$=g_0(v')[(v(\tau^-)-v')\cdot\nu]_+R^\varepsilon_\tau(\bar{\Psi})\left(1-\varepsilon-(1-\varepsilon)\hat{\rho}^\varepsilon_\tau(\bar{\Psi})-1\right)$$
$$=g_0(v')[(v(\tau^-)-v')\cdot\nu]_+R^\varepsilon_\tau(\bar{\Psi})\left(-\varepsilon-(1-\varepsilon)\hat{\rho}^\varepsilon_\tau(\bar{\Psi})\right)=R^\varepsilon_\tau(\Psi)\left(-\varepsilon-(1-\varepsilon)\hat{\rho}^\varepsilon_\tau(\bar{\Psi})\right).\tag{5.11}$$

Since the trajectory of the root particle of Ψ up to time τ is equal to the trajectory of the root particle of $\bar{\Psi}$ up to time τ and recalling that for any Φ, $\eta^\varepsilon_t(\Phi)$ is non-decreasing in t, it follows,

$$\eta^\varepsilon_\tau(\bar{\Psi})=\int_{U\times\mathbb{R}^3}g_0(\bar{v})(1-\mathbb{1}^\varepsilon_\tau[\bar{\Psi}](\bar{x},\bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v}$$
$$=\int_{U\times\mathbb{R}^3}g_0(\bar{v})(1-\mathbb{1}^\varepsilon_\tau[\Psi](\bar{x},\bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v},=\eta^\varepsilon_\tau(\Psi)\leq\eta^\varepsilon_t(\Psi),$$

and recalling (5.1),

$$C(\bar{\Psi})\leq C(\Psi).$$

Hence,

$$\rho_\tau^{\varepsilon,0}(\bar{\Psi}) = \eta_\tau^\varepsilon(\bar{\Psi})C(\bar{\Psi})\tau \leq \eta_t^\varepsilon(\Psi)C(\Psi)t = \rho_t^{\varepsilon,0}(\Psi).$$

Recalling (5.2), this implies,

$$\hat{\rho}_\tau^\varepsilon(\bar{\Psi}) = \rho_\tau^{\varepsilon,k-1}(\bar{\Psi}) \leq \rho_t^{\varepsilon,k-1}(\Psi).$$

Substituting this into (5.11),

$$\begin{aligned}\hat{P}_\tau^\varepsilon(\Psi) - R_\tau^\varepsilon(\Psi) &\geq R_\tau^\varepsilon(\Psi)\left(-\varepsilon - (1-\varepsilon)\hat{\rho}_\tau^\varepsilon(\bar{\Psi})\right)\\ &\geq R_\tau^\varepsilon(\Psi)\left(-\varepsilon - (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi)\right)\\ &= -R_\tau^\varepsilon(\Psi)\left(\varepsilon + (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi)\right).\end{aligned}$$

Recalling (5.10),

$$\begin{aligned}&\exp\left(L(\Psi)\int_\tau^t (1+2\eta_s^\varepsilon(\Psi))\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Psi) - R_\tau^\varepsilon(\Psi))\\ &\geq -\exp\left(L(\Psi)\int_\tau^t (1+2\eta_s^\varepsilon(\Psi))\,\mathrm{d}s\right)R_\tau^\varepsilon(\Psi)\left(\varepsilon + (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi)\right)\\ &= -\exp\left(L(\Psi)\int_\tau^t 2\eta_s^\varepsilon(\Psi)\,\mathrm{d}s\right)R_t^\varepsilon(\Psi)\left(\varepsilon + (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi)\right)\\ &\geq -R_t^\varepsilon(\Psi)\left(\varepsilon + (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi)\right).\end{aligned}$$

Recalling lemma 5.2 and 5.3 this gives,

$$\begin{aligned}\hat{P}_t^\varepsilon(\Psi) - R_t^\varepsilon(\Psi) &\geq \exp\left(\int_\tau^t L(\Psi)(1+2\eta_s^\varepsilon(\Psi))\,\mathrm{d}s\right)(\hat{P}_\tau^\varepsilon(\Psi) - R_\tau^\varepsilon(\Psi))\\ &\qquad + 2\eta_t^\varepsilon(\Psi)L(\Psi)R_t^\varepsilon(\Psi)\int_\tau^t \exp\left(2\eta_s^\varepsilon(\Psi)(t-s)L(\Psi)\right)\,\mathrm{d}s\\ &\geq -R_t^\varepsilon(\Psi)\left(\varepsilon + (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi)\right) - \rho_t^{\varepsilon,0}(\Psi)R_t^\varepsilon(\Psi)\\ &\geq -R_t^\varepsilon(\Psi)\left(\varepsilon + (1-\varepsilon)\rho_t^{\varepsilon,k-1}(\Psi) + \rho_t^{\varepsilon,0}(\Psi)\right)\\ &= -R_t^\varepsilon(\Psi)\rho_t^{\varepsilon,k}(\Psi) = -R_t^\varepsilon(\Psi)\hat{\rho}_t^\varepsilon(\Psi).\end{aligned}$$

This completes the proof of the inductive step which concludes the proof of the proposition. □

Proposition 5.5. *Good trees have full measure in the sense that*

$$\lim_{\varepsilon\to 0} P_t(\mathcal{MT} \setminus \mathcal{G}(\varepsilon)) = 0.$$

Proof. Firstly we prove that $\mathcal{G}(0)$ is of measure 1. To this aim recall the definition of $R(\varepsilon)$ from definition 4.2 and the definition of $\mathcal{T}_j$ (3.12). Define for $j \in \mathbb{N} \cup \{0\}$,

$$R_j(\varepsilon) := \mathcal{T}_j \cap R(\varepsilon).$$

Now note that $\mathcal{T}_0 \setminus R_0(\varepsilon)$ and $\mathcal{T}_1 \setminus R_1(\varepsilon)$ are empty, since it is not possible for a tree in $\mathcal{T}_0$ or $\mathcal{T}_1$ to have a re-collision. For a general $\Phi \in \mathcal{T}_2$, Φ is of the form,

$$\Phi = \big((x_0, v_0), (s_1, \nu_1, v_1), (s_2, \nu_2, v_2)\big).$$

If we require that $\Phi \in \mathcal{T}_2 \setminus R_2(0)$ then there must be a re-collision and since the background particles do not change velocity this requires that $v_1 = v_2$. Hence $\mathcal{T}_2 \setminus R_2(0)$ is a submanifold of $\mathcal{T}_2$ with,

$$\dim(\mathcal{T}_2 \setminus R_2(0)) < \dim(\mathcal{T}_2).$$

This argument holds true for all $j \geq 2$ and hence, because $\mathcal{MT} = \cup_{j=0}^\infty \mathcal{T}_j$ we see that,

$$P_t(\mathcal{MT} \setminus R(0)) = 1.$$

Therefore it follows, since the other requirements on $\mathcal{G}(0)$ are clear, that $P_t(\mathcal{MT} \setminus \mathcal{G}(0)) = 1$, and hence also that $P_t(\mathcal{G}(0)) = 1$.
Since $\mathcal{G}(\varepsilon)$ is increasing as ε decreases and $\lim_{\varepsilon\to 0} \mathcal{G}(\varepsilon) = \mathcal{G}(0)$ it follows by the dominated convergence theorem that,

$$\lim_{\varepsilon\to 0} P_t(\mathcal{G}(\varepsilon)) = P_t(\mathcal{G}(0)) = 1.$$

Hence

$$\lim_{\varepsilon\to 0} P_t(\mathcal{MT} \setminus \mathcal{G}(\varepsilon)) = 0,$$

as required. □

Lemma 5.6. *Recall β (4.15). For $\varepsilon > 0$, $\Phi \in \mathcal{G}(\varepsilon)$ there exists constants $C_1, C_2 > 0$ such that,*

$$\eta_t^\varepsilon(\Phi) \le C_1\varepsilon^2(\beta + V(\varepsilon)), \tag{5.12}$$

$$C(\Phi) \le C_2(\beta + V(\varepsilon)). \tag{5.13}$$

Proof. Firstly by (5.5),

$$\eta_t^\varepsilon(\Phi) = \int_{U\times\mathbb{R}^3} g_0(\bar{v})(1 - \mathbb{1}_t^\varepsilon[\Phi](\bar{x}, \bar{v}))\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{v} \le \pi\varepsilon^2 T(\beta + V(\varepsilon)).$$

So take $C_1 := \pi T$ proving (5.12). Next note,

$$\begin{aligned}\int_{\mathbb{S}^2}\int_{\mathbb{R}^3} g_0(\bar{v})[(v(t) - \bar{v})\cdot\nu]_+\,\mathrm{d}\bar{v}\,\mathrm{d}\nu &\le \pi\int_{\mathbb{R}^3} g_0(\bar{v})|v(t) - \bar{v}|\,\mathrm{d}\bar{v}\\ &\le \pi\int_{\mathbb{R}^3} g_0(\bar{v})(\mathcal{V}(\Phi) + |\bar{v}|)\,\mathrm{d}\bar{v} \le \pi(V(\varepsilon) + \beta).\end{aligned}$$

Hence by (5.1),

$$C(\Phi) \le 2\pi(V(\varepsilon) + \beta).$$

so take $C_2 := 2\pi$ which proves (5.13). □

Lemma 5.7. *For any $\delta > 0$, there exists a $\varepsilon' > 0$ such that for $0 < \varepsilon < \varepsilon'$ and for any $\Phi \in \mathcal{G}(\varepsilon)$,*

$$\hat{\rho}_t^\varepsilon(\Phi) < \delta.$$

Proof. Fix $\delta > 0$. Firstly by the above lemma,

$$\rho_t^{\varepsilon,0}(\Phi) = \eta_t(\Phi)C(\Phi)t \le C_1C_2T\varepsilon^2(\beta + V(\varepsilon))^2.$$

Recalling (4.10) there exists an $\varepsilon_1 > 0$ such that for $0 < \varepsilon < \varepsilon_1$,

$$\rho_t^{\varepsilon,0}(\Phi) < \frac{\delta}{3}.$$

Further there exists an $\varepsilon_2 > 0$ such that for $0 < \varepsilon < \varepsilon_2$,

$$\frac{1}{\sqrt{\varepsilon}}\rho_t^{\varepsilon,0}(\Phi) \le C_1C_2T\varepsilon^{3/2}(\beta + V(\varepsilon))^2 < \frac{\delta}{3}.$$

Finally there exists an $\varepsilon_3 > 0$ such that for $0 < \varepsilon < \varepsilon_3, \sqrt{\varepsilon} < \frac{\delta}{3}$. Hence take $\varepsilon' = \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3, 1\}$ then for any $\Phi \in \mathcal{G}(\varepsilon)$,

$$\begin{aligned}\hat{\rho}_t^\varepsilon(\Phi) = \rho_t^{\varepsilon,n(\Phi)}(\Phi) &= (1-\varepsilon)^{n(\Phi)}\rho_t^{\varepsilon,0}(\Phi) + (\rho_t^{\varepsilon,0}(\Phi) + \varepsilon)\sum_{j=1}^{n(\Phi)}(1-\varepsilon)^{n(\Phi)-j}\\ &\le \rho_t^{\varepsilon,0}(\Phi) + (\rho_t^{\varepsilon,0}(\Phi) + \varepsilon)\times n(\Phi) \le \rho_t^{\varepsilon,0}(\Phi) + M(\varepsilon)(\rho_t^{\varepsilon,0}(\Phi) + \varepsilon)\\ &= \rho_t^{\varepsilon,0}(\Phi) + M(\varepsilon)\rho_t^{\varepsilon,0}(\Phi) + M(\varepsilon)\varepsilon \le \rho_t^{\varepsilon,0}(\Phi) + \frac{1}{\sqrt{\varepsilon}}\rho_t^{\varepsilon,0}(\Phi) + \sqrt{\varepsilon} < \delta.\end{aligned}$$

□

Theorem 5.8. *Uniformly for $t \in [0,T]$,*

$$\lim_{\varepsilon \to 0} \|P_t - \hat{P}_t^\varepsilon\|_{TV} = 0.$$

Proof. Let $\delta > 0$ and $S \subset \mathcal{MT}$ then,

$$\begin{aligned} P_t(S) - \hat{P}_t^\varepsilon(S) &= P_t(S \cap \mathcal{G}(\varepsilon)) + P_t(S \setminus \mathcal{G}(\varepsilon)) - \hat{P}_t^\varepsilon(S \cap \mathcal{G}(\varepsilon)) - \hat{P}_t^\varepsilon(S \setminus \mathcal{G}(\varepsilon)) \\ &\leq P_t(S \cap \mathcal{G}(\varepsilon)) + P_t(S \setminus \mathcal{G}(\varepsilon)) - \hat{P}_t^\varepsilon(S \cap \mathcal{G}(\varepsilon)). \end{aligned}$$

By proposition 5.5 for ε sufficiently small,

$$P_t(S \setminus \mathcal{G}(\varepsilon)) \leq P_t(\mathcal{MT} \setminus \mathcal{G}(\varepsilon)) < \frac{\delta}{3}.$$

Hence,

$$P_t(S) - \hat{P}_t^\varepsilon(S) < P_t(S \cap \mathcal{G}(\varepsilon)) - \hat{P}_t^\varepsilon(S \cap \mathcal{G}(\varepsilon)) + \frac{\delta}{3}. \tag{5.14}$$

Recall the definition of $\zeta(\varepsilon)$ in (4.3). It is clear that this implies

$$\zeta(\varepsilon) \leq 1. \tag{5.15}$$

Hence by the above lemma, for ε sufficiently small and $\Phi \in \mathcal{G}(\varepsilon)$,

$$\zeta(\varepsilon)\hat{\rho}_t^\varepsilon(\Phi) \leq \hat{\rho}_t^\varepsilon(\Phi) < \frac{\delta}{3}.$$

Further by proposition 5.1, for $\Phi \in \mathcal{G}(\varepsilon)$,

$$-\hat{P}_t^\varepsilon(\Phi) \leq -R_t^\varepsilon(\Phi) + \hat{\rho}_t^\varepsilon(\Phi) R_t^\varepsilon(\Phi). \tag{5.16}$$

The Binomial inequality states that for $x \geq -1$ and $N \in \mathbb{N}$,

$$(1+x)^N \geq 1 + Nx.$$

Hence for $\varepsilon > 0$ such that $\frac{4}{3}\pi\varepsilon^3 \leq 1$ we apply this to $\zeta(\varepsilon)$ recalling (2.1),

$$\zeta(\varepsilon) = (1 - \frac{4}{3}\pi\varepsilon^3)^N \geq 1 - N\frac{4}{3}\pi\varepsilon^3 = 1 - \frac{4}{3}\pi\varepsilon. \tag{5.17}$$

Hence for ε sufficiently small (5.16) gives,

$$\begin{aligned} P_t(\Phi) - \hat{P}_t^\varepsilon(\Phi) &\leq P_t(\Phi) - R_t^\varepsilon(\Phi) + \hat{\rho}_t^\varepsilon(\Phi) R_t^\varepsilon(\Phi) \\ &= P_t(\Phi) - \zeta(\varepsilon)P_t(\Phi) + \hat{\rho}_t^\varepsilon(\Phi)\zeta(\varepsilon)P_t(\Phi) \leq \frac{4}{3}\pi\varepsilon P_t(\Phi) + \frac{\delta}{3}P_t(\Phi) \leq \frac{2\delta}{3}P_t(\Phi). \end{aligned}$$

This holds for all $\Phi \in \mathcal{G}(\varepsilon)$ with ε sufficiently small hence

$$P_t(S \cap \mathcal{G}(\varepsilon)) - \hat{P}_t^\varepsilon(S \cap \mathcal{G}(\varepsilon)) \leq \frac{2\delta}{3}P_t(S \cap \mathcal{G}(\varepsilon)) \leq \frac{2\delta}{3}. \tag{5.18}$$

By substituting (5.18) into (5.14), for ε sufficiently small,

$$P_t(S) - \hat{P}_t^\varepsilon(S) < \delta. \tag{5.19}$$

Since ε did not depend on S this holds true for every $S \subset \mathcal{MT}$. Hence for any $S \subset \mathcal{MT}$,

$$\hat{P}_t^\varepsilon(S) - P_t(S) = (1 - \hat{P}_t^\varepsilon(\mathcal{MT} \setminus S)) - (1 - P_t(\mathcal{MT} \setminus S)) \quad = P_t(\mathcal{MT} \setminus S) - \hat{P}_t^\varepsilon(\mathcal{MT} \setminus S) < \delta.$$

This together with (5.19) gives that, for ε sufficiently small, for any $S \subset \mathcal{MT}$,

$$|P_t(S) - \hat{P}_t^\varepsilon(S)| < \delta,$$

which completes the proof of the theorem. □

We can now prove the main theorem of the paper, theorem 2.3, which follows from the above theorem.

Proof of theorem 2.3. Recall for $\Omega \subset U \times \mathbb{R}^3$,

$$S_t(\Omega) := \{\Phi \in \mathcal{MT} : (x(t), v(t)) \in \Omega\}.$$

By theorem 3.1,

$$\int_\Omega f_t(x,v)\,\mathrm{d}x\,\mathrm{d}v = \int_{S_t(\Omega)} P_t(\Phi)\,\mathrm{d}\Phi = P_t(S_t(\Omega)).$$

Also by definition of $\hat{P}_t^\varepsilon$,

$$\int_\Omega \hat{f}_t^N(x,v)\,\mathrm{d}x\,\mathrm{d}v = \int_{S_t(\Omega)} \hat{P}_t^\varepsilon(\Phi)\,\mathrm{d}\Phi = \hat{P}_t^\varepsilon(S_t(\Omega)).$$

Hence by theorem 5.8,

$$\lim_{N\to\infty} \sup_{\Omega\subset U\times\mathbb{R}^3} |\int_\Omega \hat{f}_t^N(x,v) - f_t(x,v)\,\mathrm{d}x\,\mathrm{d}v| = \lim_{\varepsilon\to 0} \sup_{\Omega\subset U\times\mathbb{R}^3} |P_t(S_t(\Omega)) - \hat{P}_t^\varepsilon(S_t(\Omega))| = 0,$$

which completes the proof. □

6. Proof of auxiliary results

6.1. **Particle dynamics.**

Proof of Prop 2.1. The dynamics become undefined if there is instantaneously more than one background particle colliding with the tagged particle or if the tagged particle experiences an infinite number of collisions in finite time. We adapt a similar proof for the full hard-spheres dynamics from [14, Prop 4.1.1].

Let $R > 0$ and $\delta < \varepsilon/2$ such that there exists a $K \in \mathbb{N}$ with $T = K\delta$. Denote the ball of radius R about x in $\mathbb{R}^3$ by $B_R(x)$. For the initial position of the tagged particle $(x_0, v_0) \in U \times B_R(0)$ fixed define $I(x_0, v_0) \subset (U \times \mathbb{R}^3)^N$ by,

$$\begin{aligned} I(x_0,v_0) := \{(x_1,v_1),\dots,(x_N,v_N) \in (U\times B_R(0))^N :\ &\text{the tagged particle collides} \\ &\text{with at least two background particles in the time interval } [0,\delta]\}. \end{aligned}$$

We bound the volume of this set. Firstly define

$$I^1(x_0,v_0) := \{(x_1,v_1) \in U \times B_R(0) : \varepsilon \le |x_0 - x_1| \le \varepsilon + 2R\delta\}.$$

It can be seen that for some C,

$$|I^1(x_0,v_0)| \le CR^3 \times (2R\delta)^3.$$

Since $I(x_0, v_0)$ is a subset of,

$$\begin{aligned} \{(x_1,v_1),\dots,(x_N,v_N) \in (U\times B_R(0))^N : \exists\, 1 \le i < j \le N \text{ such that} \\ \varepsilon \le |x_0 - x_i| \le \varepsilon + 2R\delta \text{ and } \varepsilon \le |x_0 - x_j| \le \varepsilon + 2R\delta\} \end{aligned}$$

The above estimate gives, for some constant $C = C(N, \varepsilon)$,

$$\begin{aligned} |I(x_0,v_0)| &\le CR^{3(N-2)} \times (R^6\delta^3)^2 \\ &\le CR^{3(N+2)}\delta^6. \end{aligned}$$

Hence if we define,

$$I := \cup\{I(x_0,v_0) : (x_0,v_0) \in U \times B_R(0)\},$$

it follows,

$$|I| \le CR^{3(N+3)}\delta^6.$$

Hence there exists a subset $I_0(\delta, R)$ of measure at most $CR^{3(N+3)}\delta^6$ such that for any initial configuration in $(U \times B_R(0))^{N+1} \setminus I_0(\delta, R)$ the tagged particle experiences at most one collision in $[0, \delta]$.

Now consider the system at time δ. Since all particles had initial velocity in $B_R(0)$ and the tagged particle had at most one collision in time $[0, \delta]$ the velocity of the tagged particle at time δ is in $B_{2R}(0)$. By the same arguments above there exists a set $I_1(\delta, R)$ of measure at most

$CR^{3(N+3)}\delta^6$ for some new constant C such that for any initial configuration in $(U \times B_R(0))^{N+1} \setminus I_0(\delta, R) \cup I_1(\delta, R)$ the tagged particle experiences at most one collision in $[0, \delta]$ and at most one collision in $[\delta, 2\delta]$ and thus the dynamics are well defined up to 2δ.
Continue this process K times defining the set,

$$I(\delta, R) := \cup_{j=0}^{K-1} I_j(\delta, R),$$

which has measure at most $CR^{3(N+3)}\delta^6$ for some new constant C and such that for any initial configuration in $(U \times B_R(0))^{N+1} \setminus I(\delta, R)$ the tagged particle has at most one collision per time interval $[j\delta, (j+1)\delta]$ and hence the dynamics are well defined up to time T. Defining,

$$I(T, R) := \cap_{\delta > 0} I(\delta, R),$$

if follows $I(T, R)$ is of measure zero and for any any initial configuration in $(U \times B_R(0))^{N+1} \setminus I(T, R)$ the dynamics are well defined up to time T.
Finally take,

$$I := \cup_{R \in \mathbb{N}} I(T, R)$$

and note that I is a countable union of measure zero sets and for any initial configuration in $(U \times \mathbb{R}^3) \setminus I$ the dynamics are well defined up to time T. □

6.2. **Semigroup theory.**

Lemma 6.1. *A is the generator of the substochastic semigroup (c.f.* [3, Section 10.2]*)* $T(t) : L^1(U \times \mathbb{R}^3) \to L^1(U \times \mathbb{R}^3)$ *given by,*

$$(T(t)f)(x, v) := \exp\left(-t \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu\right) f(x - tv, v). \tag{6.1}$$

Proof. We seek to apply theorem 10.4 of [3]. Conditions (A1), (A2) trivially hold since $F = 0$ in our situation. As for (A3),

$$\nu(x, v) = \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu. \tag{6.2}$$

This is locally integrable so (A3) holds. Hence we can apply the theorem. In our case,

$$\varphi(x, v, t, s) = x - (t + s)v.$$

So the semigroup is given by

$$(T(t)f)(x, v) := \exp\left(-t \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu\right) f(x - tv, v),$$

as required. □

Remark 6.2. *For* $v, v_* \in \mathbb{R}^3, v \neq v_*$ *define the Boltzmann kernel* k *by,*

$$k(v, v_*) := \frac{1}{|v - v_*|} \int_{E_{v,v_*}} g_0(w) \, \mathrm{d}w, \tag{6.3}$$

where $E_{v,v_*} = \{w \in \mathbb{R}^3 : w \cdot (v - v_*) = v \cdot (v - v_*)\}$*. By the use of Carleman's representation, first described in* [9] *(see also* [4, Section 3]*),*

$$\begin{aligned}(Bf)(x, v) &= Q^+[f](x, v) \\ &= \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} f(x, v') g_0(\bar{v}')[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu \\ &= \int_{\mathbb{R}^3} k(v, v_*) f(x, v_*) \, \mathrm{d}v_*. \end{aligned} \tag{6.4}$$

Lemma 6.3. *There exists a* $C > 0$ *such that for any* $V > 0$*,*

$$\int_{|v| > V} k(v, v_*) \, \mathrm{d}v \leq C, \tag{6.5}$$

for almost all $|v_*| \leq V$*.*

Proof. A similar calculation to the proof of theorem 2.1 in [2] (see also [27]) allows us to show that in our case, for $v \neq v_*$,

$$k(v, v_*) \leq \frac{2\pi}{|v - v_*|} h\left(\frac{1}{4}\left[|v - v_*| + \frac{|v|^2 - |v_*|^2}{|v - v_*|}\right]^2\right).$$

Where $h : [0, \infty) \to \mathbb{R}$ is given by,

$$h(r) := \int_r^\infty z \sup_{\{v \in \mathbb{R}^3 : |v| = z\}} g_0(v) \, \mathrm{d}z.$$

Thus we can follow the calculations in [3, Ex 10.29] to see that, for $d = 3$, for almost all $|v^*| \leq V$ and using assumption (2.4),

$$\int_{|v|>V} k(v, v_*) \, \mathrm{d}v \leq C\Gamma < \infty,$$

where $C > 0$ is some geometric constant and

$$\Gamma := \int_0^\infty r h(r^2) \, \mathrm{d}r.$$

□

Proof of proposition 3.5. By the definition of A and B, (2.5) is now,

$$\begin{cases} \partial_t f_t(x, v) & = (Af + Bf)(x, v) \\ f_{t=0}(x, v) & = f_0(x, v). \end{cases} \tag{6.6}$$

By lemma 6.1 and equations (10.86), (10.88) of [3], [3, corollary 5.17] holds for our equation, so we apply [3, theorem 5.2]. Lemma 6.3 allow us to further apply [3, theorem 10.28]. Therefore we have an honest C_0 semigroup of contractions generated by $\overline{A + B}$, which we denote by $G(t)$. Finally by [1, theorem 3.1.12], (6.6) has a unique mild solution for each $f_0 \in L^1(U \times \mathbb{R}^3)$.
It remains to show (3.8) and (3.9). Firstly we find a bound for the operator B. Then we prove (3.8) for $f_0 \in D(A)$ first before generalising to all f_0. Recall (4.15). As g_0 is normalized,

$$\begin{aligned} \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu &\leq \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})|v - \bar{v}| \, \mathrm{d}\bar{v} \, \mathrm{d}\nu \\ &\leq \pi \int_{\mathbb{R}^3} g_0(\bar{v})(|v| + |\bar{v}|) \, \mathrm{d}\bar{v} \, \mathrm{d}\nu \leq \pi(|v| + \beta). \end{aligned} \tag{6.7}$$

By (10.6) in [3] for $f \in L^1(U \times \mathbb{R}^3)$ and by (6.7) recalling (6.2),

$$\begin{aligned} \int_{U \times \mathbb{R}^3} Bf(x, v) \, \mathrm{d}x \, \mathrm{d}v &= \int_{U \times \mathbb{R}^3} \int_{\mathbb{R}^3} k(v, v_*) f(x, v_*) \, \mathrm{d}v_* \, \mathrm{d}x \, \mathrm{d}v \\ &= \int_{U \times \mathbb{R}^3} \nu(x, v) f(x, v) \, \mathrm{d}x \, \mathrm{d}v \\ &= \int_{U \times \mathbb{R}^3} f(x, v) \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} g_0(\bar{v})[(v - \bar{v}) \cdot \nu]_+ \, \mathrm{d}\bar{v} \, \mathrm{d}\nu \, \mathrm{d}x \, \mathrm{d}v \\ &\leq \int_{U \times \mathbb{R}^3} f(x, v) \pi(|v| + \beta) \, \mathrm{d}x \, \mathrm{d}v. \end{aligned} \tag{6.8}$$

Now suppose that $f_0 \in D(A)$. Then [3, corollary 5.17] holds and so we apply [3, corollary 5.8] which gives,

$$f_t = T(t) f_0 + \int_0^t G(t - \theta) B T(\theta) f_0 \, \mathrm{d}\theta. \tag{6.9}$$

Where $G(t)$ is the contraction semigroup generated by $\overline{A+B}$ and $T(t)$ the contraction semigroup generated by A. Hence by (6.8),

$$\begin{aligned}\int_{U\times\mathbb{R}^3} & f_t(x,v)(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \\ &\leq \int_{U\times\mathbb{R}^3} T(t)f_0(x,v)(1+|v|) + \int_0^t G(t-\theta)BT(\theta)f_0\,\mathrm{d}\theta(x,v)(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \\ &\leq \int_{U\times\mathbb{R}^3} f_0(x,v)(1+|v|) + tf_0(x,v)(1+|v|)\pi(|v|+\beta)\,\mathrm{d}x\,\mathrm{d}v.\end{aligned}$$

Noting that $t\in[0,T]$ and recalling our assumption on f_0 in (2.2), this is bounded.
Consider a general f_0 in (2.2), not necessarily in $D(A)$. Suppose for contradiction that (3.8) is not true. Hence there exists a $t\in[0,T]$ such that for any $C>0$ there exists an $R>0$ such that,

$$\int_U\int_{B_R(0)} f_t(x,v)(1+|v|)\,\mathrm{d}v\,\mathrm{d}x \geq C.$$

$D(A)$ is dense in $L^1(U\times\mathbb{R}^3)$ because it contains, for example, smooth compactly supported functions. Hence for any $n\in\mathbb{N}$ there exists an $f_0^n\in L^1(U\times\mathbb{R}^3)$ such that $f_0^n\in D(A)$, $f_0^n\geq 0$, there exists a $C_1>0$ such that

$$\int_{U\times\mathbb{R}^3} f_0^n(x,v)(1+|v|^2)\,\mathrm{d}x\,\mathrm{d}v \leq C_1, \tag{6.10}$$

and that,

$$\int_{U\times\mathbb{R}^3} |f_0^n(x,v)-f_0(x,v)|(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \leq \frac{1}{n}. \tag{6.11}$$

Now define f_t^n to be the solution of (2.5) with initial data given by f_0^n. In this case (6.9) gives,

$$f_t^n = T(t)f_0^n + \int_0^t G(t-\theta)BT(\theta)f_0^n\,\mathrm{d}\theta.$$

The argument above for $f_0^n\in D(A)$ together with (6.10) gives that there exists a $C_2>0$ independent of n such that,

$$\int_{U\times\mathbb{R}^3} f_t^n(x,v)(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \leq C_2.$$

By our contradiction assumption there exists an $R>0$ such that,

$$\int_U\int_{B_R(0)} f_t(x,v)(1+|v|)\,\mathrm{d}v\,\mathrm{d}x \geq 2C_2.$$

These two bounds together give that,

$$\begin{aligned}\int_{U\times\mathbb{R}^3} & |f_t(x,v)-f_t^n(x,v)|(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \\ &\geq \int_U\int_{B_R(0)} |f_t(x,v)-f_t^n(x,v)|(1+|v|)\,\mathrm{d}v\,\mathrm{d}x \\ &\geq \left|\int_U\int_{B_R(0)} f_t(x,v)(1+|v|)\,\mathrm{d}v\,\mathrm{d}x - \int_U\int_{B_R(0)} f_t^n(x,v)(1+|v|)\,\mathrm{d}v\,\mathrm{d}x\right| \geq C_2. \qquad (6.12)\end{aligned}$$

However since $G(t)$ is a contraction semigroup it follows for $n>1/C_2$ by (6.11),

$$\begin{aligned}\int_{U\times\mathbb{R}^3} (f_t(x,v)-f_t^n(x,v))(1+|v|)\,\mathrm{d}x\,\mathrm{d}v &= \int_{U\times\mathbb{R}^3} G(t)(f_0-f_0^n)(x,v)(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \\ &\leq \int_{U\times\mathbb{R}^3} (f_0(x,v)-f_0^n(x,v))(1+|v|)\,\mathrm{d}x\,\mathrm{d}v \leq \frac{1}{n} < C_2.\end{aligned}$$

By the same argument,

$$\int_{U\times\mathbb{R}^3} (f_t^n(x,v) - f_t(x,v))(1+|v|)\,\mathrm{d}x\,\mathrm{d}v < C_2.$$

Hence we have a contradiction with (6.12) which completes the proof of (3.8).
To show (3.9) fix $t \in [0,T]$. By (6.8),

$$\int_{U\times\mathbb{R}^3} Bf_t(x,v)\,\mathrm{d}x\,\mathrm{d}v \le \int_{U\times\mathbb{R}^3} f_t(x,v)\pi(|v|+\beta)\,\mathrm{d}x\,\mathrm{d}v.$$

By the above calculations f_t has finite first moment so this is finite as required. □

References

[1] Wolfgang Arendt, Charles J. K. Batty, Matthias Hieber, and Frank Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser/Springer Basel AG, Basel, second edition, 2011.

[2] Luisa Arlotti and Bertrand Lods. Integral representation of the linear Boltzmann operator for granular gas dynamics with applications. *Journal of Statistical Physics*, 129(3):517–536, 2007.

[3] Jacek Banasiak and Luisa Arlotti. *Perturbations of positive semigroups with applications*. Springer Monographs in Mathematics. Springer-Verlag London, Ltd., London, 2006.

[4] Marzia Bisi, José A Cañizo, and Bertrand Lods. Entropy dissipation estimates for the linear Boltzmann operator. *Journal of Functional Analysis*, 269(4):1028—1069, 2015.

[5] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Inventiones mathematicae*, 203(2):493–553, 2015.

[6] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. From hard spheres dynamics to the Stokes–Fourier equations: an L2 analysis of the Boltzmann–Grad limit. *Comptes Rendus Mathematique*, 353(7):623–627, 2015.

[7] Carlo Boldrighini, Leonid A. Bunimovich, and Yakov. G. Sinaĭ. On the Boltzmann equation for the Lorentz gas. *J. Statist. Phys.*, 32(3):477–501, 1983.

[8] Max Born and HS Green. A general kinetic theory of liquids. i. the molecular distribution functions. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 188, pages 10–18. The Royal Society, 1946.

[9] Torsten Carleman. *Problemes mathématiques dans la théorie cinétique de gaz*, volume 2. Almqvist & Wiksells boktr, 1957.

[10] Carlo Cercignani. *The Boltzmann equation and its applications*, volume 67 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1988.

[11] Carlo Cercignani, Reinhard Illner, and Mario Pulvirenti. *The mathematical theory of dilute gases*, volume 106 of *Applied Mathematical Sciences*. Springer-Verlag, 1994.

[12] Donald L. Cohn. *Measure theory*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, New York, second edition, 2013.

[13] L. Desvillettes and V. Ricci. The Boltzmann-Grad limit of a stochastic Lorentz gas in a force field. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 2(2):637–648, 2007.

[14] Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European mathematical society, 2013.

[15] Giovanni Gallavotti. *Statistical Mechanics. A Short Treatise*. Theoretical and Mathematical Physics. Springer-Verlag Berlin Heidelberg, 1999.

[16] François Golse. On the periodic Lorentz gas and the Lorentz kinetic equation. *Ann. Fac. Sci. Toulouse Math. (6)*, 17(4):735–749, 2008.

[17] Oscar E. Lanford. *Dynamical Systems, Theory and Applications: Battelle Seattle 1974 Rencontres*, chapter Time evolution of large classical systems, pages 1–111. Springer Berlin Heidelberg, Berlin, Heidelberg, 1975.

[18] Joel L. Lebowitz and Herbert Spohn. Transport properties of the Lorentz gas: Fourier's law. *J. Statist. Phys.*, 19(6):633–654, 1978.

[19] Joel L. Lebowitz and Herbert Spohn. Microscopic basis for Fick's law for self-diffusion. *J. Statist. Phys.*, 28(3):539–556, 1982.

[20] Joel L. Lebowitz and Herbert Spohn. Steady state self-diffusion at low density. *J. Statist. Phys.*, 29(1):39–55, 1982.

[21] Hendrik Lorentz. The motion of electrons in metallic bodies i. *Koninklijke Nederlandsche Akademie van Wetenschappen Proceedings*, 7:438–453, 1905.

[22] Jens Marklof. Kinetic transport in crystals. In *XVIth International Congress on Mathematical Physics*, pages 162–179. World Sci. Publ., Hackensack, NJ, 2010.

[23] Jens Marklof and Andreas Strömbergsson. The Boltzmann-Grad limit of the periodic Lorentz gas. *Ann. of Math. (2)*, 174(1):225–298, 2011.
[24] Karsten Matthies and Florian Theil. Validity and non-validity of propagation of chaos. *Analysis and Stochastics of growth processes and Interface Models*, pages 101–119, 2008.
[25] Karsten Matthies and Florian Theil. Validity and failure of the Boltzmann approximation of kinetic annihilation. *Journal of nonlinear science*, 20(1):1–46, 2010.
[26] Karsten Matthies and Florian Theil. A semigroup approach to the justification of kinetic theory. *SIAM Journal on Mathematical Analysis*, 44(6):4345–4379, 2012.
[27] Frederic A Molinet. Existence, uniqueness and properties of the solutions of the Boltzmann kinetic equation for a weakly ionized gas. i. *Journal of Mathematical Physics*, 18(5):984–996, 1977.
[28] Amnon Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, 1983.
[29] Mario Pulvirenti, Chiara Saffirio, and Sergio Simonella. On the validity of the Boltzmann equation for short range potentials. *Reviews in Mathematical Physics*, 26(02), 2014.
[30] Herbert Spohn. The Lorentz process converges to a random flight process. *Comm. Math. Phys.*, 60(3):277–290, 1978.
[31] Herbert. Spohn. Kinetic equations from Hamiltonian dynamics: the Markovian approximations. In *Kinetic theory and gas dynamics*, volume 293 of *CISM Courses and Lect.*, pages 183–211. Springer, Vienna, 1988.
[32] Herbert Spohn. *Large Scale Dynamics of Interacting Particles*. Texts and Monographs in Physics. Springer Berlin Heidelberg, 1991.
[33] Kōhei Uchiyama. Derivation of the Boltzmann equation from particle dynamics. *Hiroshima Math. J.*, 18(2):245–297, 1988.
[34] Henk van Beijeren, Oscar E. Lanford, III, Joel L. Lebowitz, and Herbert Spohn. Equilibrium time correlation functions in the low-density limit. *J. Statist. Phys.*, 22(2):237–257, 1980.

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom
E-mail address: K.Matthies@bath.ac.uk

Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom
E-mail address: G.Stone@bath.ac.uk

Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom
E-mail address: f.theil@warwick.ac.uk