Kac’s Process with Hard Potentials and a Moderate Angular Singularity

DANIEL HEYDECKER

Communicated by C. MOUHOT

Abstract

We investigate Kac’s many-particle stochastic model of gas dynamics in the case of hard potentials with a moderate angular singularity, and show that the noncutoff particle system can be obtained as the limit of cutoff systems, with a rate independent of the number of particles $N$. As consequences, we obtain a wellposedness result for the corresponding Boltzmann equation and propagation of chaos in the many-particle limit $N \to \infty$.

1. Introduction and Main Results

Let us consider Kac’s model [26] for the behaviour of a dilute gas. We consider an ensemble of $N$ indistinguishable particles, with velocities $V^1_t, \ldots, V^N_t \in \mathbb{R}^d$, which are encoded in the empirical velocity distribution $\mu^N_t = N^{-1} \sum_{i=1}^N \delta_{V^i_t}$. The rates of each possible collision are governed by a collision kernel $B : \mathbb{R}^d \times S^{d-1} \to [0, \infty)$, which reflects the physics of the underlying system, and the dynamics can be described informally as follows. For every unordered pair of particles with velocities $v, v_*$, the velocities change to

$$v \mapsto v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2\sigma}; \quad v_* \mapsto v_*' = \frac{v + v_*}{2} - \frac{|v - v_*|}{2\sigma} \quad (1)$$

at a rate $2B(v - v_*, \sigma)d\sigma/N$. These collisions preserve energy and momentum, so that the total energy $\langle |v|^2, \mu \rangle$ and momentum $\langle v, \mu \rangle$ are preserved as time runs. Let us define the Boltzmann Sphere $S$ as those measures on $\mathbb{R}^d$ with normalised mass, momentum and energy:

$$\langle 1, \mu \rangle = 1; \quad \langle v, \mu \rangle = 0; \quad \langle |v|^2, \mu \rangle = 1.$$ (2)
We also write $\mathcal{S}_N$ for the subspace of $\mathcal{S}$ consisting of normalised empirical measures on $N$ points; we equip both spaces with the (Monge-Kantorovich) Wasserstein distance $w_1$ which is recalled in (13), below. Due to Galilean invariance, the collision kernel is assumed to be of the form $B(v, \sigma) = B(|v|, \cos \theta)$, where $\theta$ is the deflection angle given by $\cos \theta = \sigma \cdot v/|v|$. We will consider the case of noncutoff hard potentials, where the kernel has the form

$$B(v, \sigma) = |v|^\gamma (\sin \theta)^{2-d} \beta(\theta); \quad \beta(\theta) \sim \theta^{1-v} \quad \text{as } \theta \downarrow 0$$

with $\gamma \in [0, 1], v \in [0, 1)$. We will also assume that $\beta$ takes the form $\beta(\theta) = b(\cos \theta)$, for a convex function $b : (-1, 1) \to [0, \infty)$; (3) then rearranges to $b(x) \sim (1-x)^{-1/2-v/2}$ as $x \uparrow 1$. Thanks to the symmetry of collisions, we may assume further that $b$ is supported on $[0, 1)$; see the discussion in Alexandre et al. [2].

In $d = 3$, such kernels arise when modelling particles interacting through a repulsive potential $V(r) = r^{-s}, s > 5$, with $\gamma = \frac{s-5}{s-1}, v = \frac{2}{s-2}$. Importantly, the kernel is not integrable, due to the abundance of grazing collisions, reflected in the non-integrable singularity of $\beta$ as $\theta \downarrow 0$. The cases we consider have a mild angular singularity, so that

$$\int_{\mathbb{S}^{d-1}} B(v, \sigma) d\sigma = \infty; \quad \int_{\mathbb{S}^{d-1}} \theta B(v, \sigma) d\sigma < \infty.$$  

The divergence of the total rate $\int B d\sigma$ implies that each pair of particles undergoes infinitely many collisions on any nontrivial time interval, and there is work to be done in understanding the informal description of the dynamics above. Formally, a Kac process is a Markov process $\mu^N_t$ in $(\mathcal{S}_N, w_1)$, with càdlàg paths and generator, defined for Lipschitz $F : (\mathcal{S}_N, w_1) \to \mathbb{R}$:

$$(G^N F)(\mu^N) = N \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}^{d-1}} (F(\mu^N, v, v', \sigma) - F(\mu^N))$$

$$\times B(v - v', \sigma) \mu^N(dv') \mu^N(dv) d\sigma.$$  

Here we have written $\mu^N, v, v', \sigma = \mu^N + N^{-1}(\delta_{v'} + \delta_{v'} - \delta_v - \delta_{v'})$ for the measure $\mu^N$ replacing precollisional velocities $v, v' \in \text{supp}(\mu^N)$ by postcollisional velocities $v', v'$, given in terms of $v, v, \sigma$ by (1). One can check that $w_1(\mu^N, v, v', \sigma, \mu^N) \leq 2|v - v'|/N$, and so, thanks to (4), the integral written above is convergent for Lipschitz $F$. However, the total rate is still infinite, and it is not a priori clear that the associated martingale problem is well-posed.

**Labelled Versus Unlabelled Dynamics** Let us briefly mention that it is also possible to work with a labelled Kac process $\mathcal{V}^N_i = (V^N_{t_1}, \ldots, V^N_{t_N}) \in (\mathbb{R}^d)^N$, where each particle is assigned a label $i = 1, \ldots, N$; in this case, the same normalisation (2) now defines a subspace $\mathcal{S}_N \subset (\mathbb{R}^d)^N$. The interchangeability of the particles means that such processes have a Sym($N$)-symmetry by exchanging

---

1 Here, and throughout, angle brackets $\langle \cdot, \cdot \rangle$ denote integration against a measure, and $\nu$ denotes the identity function on $\mathbb{R}^d$. 
the labels; we identify the orbit of $\mathcal{V}^N \in \mathbb{S}_N$ with its empirical measures, and $\mathbb{S}_N \equiv (\mathbb{R}^d)^N / \text{Sym}(N)$. Let us also write $\theta^N : \mathbb{S}_N \to \mathbb{S}_N$ for the map taking a $N$ velocities $\mathcal{V}^N = (V^1, \ldots, V^N) \in \mathbb{S}_N$ to the associated empirical measure $\mu^N = N^{-1} \sum_i \delta_{V_i}$. 

Grad’s Angular Cutoff A convenient simplification of the dynamics above is Grad’s angular cutoff, which truncates $B$ near small values of $\theta$; one might hope that this truncation preserves, in some meaningful sense, the physics of the system under consideration. Let us define, for $K > 0$, 

$$B_K(v, \sigma) = B(v, \sigma)1(\theta > \theta_0(K)),$$

(6)

where $\theta_0(K)$ is chosen so that $\int_{S^{d-1}} B_K(v, \theta) d\sigma = K |v|^\gamma$. We can now consider the $K$-cutoff Kac processes $\mathcal{V}^N_t$, $K \geq 0$, on these kernels, with generator defined analogously to (5). In this case, the total rate is finite, and the associated martingale problem has uniqueness in law. The key intermediate result of this paper, summarised in Lemma 8 and Corollary 1, is that any noncutoff Kac process $\mu^N_t$ on $N$ particles can be obtained as the limit of cutoff process $\mu^N_{t,K}$ of cutoff processes as $K \to \infty$, in an $N$-uniform way.

Measure Solutions to the Boltzmann Equation Kac introduced the stochastic system described above in an effort to justify the spatially homogeneous Boltzmann Equation; following previous works [17,33,34], we will consider measure-valued solutions. For a measure $\mu$ with finite second moment, we define the Boltzmann collision operator by specifying, for all Lipschitz $f : \mathbb{R}^d \to \mathbb{R}$,

$$\langle f, Q(\mu) \rangle = \int_{\mathbb{R}^d \times \mathbb{R}^d \times S^{d-1}} \left\{ f(v') + f(v_*) - f(v) - f(v_*) \right\} \times B(v - v_*, d\sigma) \mu(dv) \mu(dv_*).$$

(7)

The same argument as (5) below shows that this integral is well-defined for Lipschitz $f$, but $Q$ must be interpreted as a distribution, for example in the negative Sobolev space $W^{-1,\infty}(\mathbb{R}^d) = W^{1,\infty}(\mathbb{R}^d)^*$, rather than a signed measure. We say that a family $(\mu_t)_{t \geq 0}$ of measures in $\mathcal{S}$ satisfies the Boltzmann equation if, for any Lipschitz $f$ of compact support,

$$\forall t \geq 0 \quad \langle f, \mu_t \rangle = \langle f, \mu_0 \rangle + \int_0^t \langle f, Q(\mu_s) \rangle ds.$$  

(BE)

Replacing $Q$ with the cutoff collision operator $Q_K$, with $B_K$ in place of $B$, we obtain the corresponding $K$-cutoff Boltzmann equations:

$$\mu_t = \mu_0 + \int_0^t Q_K(\mu_s) ds.$$  

(BE_K)

This may be interpreted in a similar way to (BE), above. It is known [30, Theorem 1.5d] that (BE_K) admits unique energy-conserving solutions for any initial data $\mu_0 \in \mathcal{S}$, which we write as $\phi^K_t(\mu_0)$, $t \geq 0$. The main contribution of this paper to the study of the Boltzmann equation follows the coupling argument, viewing the Kac process as a discrete approximation to the Boltzmann equation. The $N$-uniformity of our results at the level of the Kac process allows us to transfer the
stability in the initial data, and that the cutoff system converges to the noncutoff system as the cutoff is removed, from the Kac process to the Boltzmann equation, leading to Theorem 1 below.

**Notation Regarding Moments** Our estimates will frequently include moments of the Kac process or Boltzmann flow, and it is convenient to introduce notation to deal with this. Let us define $S^k$ as those measures $\mu \in S$ with a finite $k$th moment $\langle |v|^k, \mu \rangle$, and define

$$\Lambda_k(\mu) := \langle (1 + |v|^{2k/2}, \mu); \quad \Lambda_k(\mu, \nu) := \max(\Lambda_k(\mu), \Lambda_k(\nu)). \quad (8)$$

With this notation, we define

$$S^k_a := \{\mu \in S : \Lambda_k(\mu) \leq a\}. \quad (9)$$

**An Optimal Transportation Cost** The key novelty of the current work is to take advantage of a cancellation which occurs when working with a tailor-made optimal transportation cost on measures. We write, for $p \geq 0$,

$$d_p(v, w) = (1 + |v|^p + |w|^p)^{1/2} |v - w|. \quad (10)$$

We now define the optimal transport cost $T_p$ on probability measures with $p + 2$ moments as

$$T_p(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} d_p(v, w)^2 \pi(\, dv, dw) \right)^{1/2}, \quad (11)$$

where $\Pi(\mu, \nu)$ denotes the set of couplings of $\mu$ and $\nu$; we emphasise that $p$ here denotes a weighting, and not an exponent in the usual sense of Wasserstein $p$ metrics. These optimal transport costs do not define metrics, since the triangle inequality fails, but instead are semimetrics in that the triangle inequality is replaced by

$$T_p(\mu, \nu) \leq C \left( T_p(\mu, \xi) + T_p(\xi, \nu) \right) \quad (12)$$

for some $C = C(p)$ and all $\mu, \nu, \xi \in S^{p+2}$. We will occasionally make use of the Wasserstein$_1$ metric

$$w_1(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |v - w| \pi(\, dv, dw), \quad (13)$$

and note the easy interpolation estimate, for any $p' > p + 2$ and any $\mu, \nu \in S^{p'}$,

$$w_1(\mu, \nu) \leq T_p(\mu, \nu) \leq w_1(\mu, \nu)^{\alpha} \Lambda_{p'}(\mu, \nu) \quad (14)$$

for some $\alpha = \alpha(p, p') > 0$. Although we will not make use of this fact, it is straightforward to verify that the optimal transport costs $T_p$ defined in this way are equivalent to the usual Wasserstein metric $w_{p+2}$ of order $p + 2$, in the sense that they produce the same convergent sequences: for any $\mu \in S^{p+2}$ and sequence $(\mu^n) \subset S^{p+2}$, it holds that $T_p(\mu^n, \mu) \to 0$ if and only if $w_{p+2}(\mu^n, \mu) \to 0$. We refer the interested reader to [25, Section 2.1, pp. 58–64] for details.
1.1. Main results

With this notation and terminology fixed, we can now state our main results. Our first result concerns the wellposedness and stability of the limit equation (BE).

**Theorem 1.** (Uniqueness and Stability) Let $B$ be a kernel of the form (3) with $\gamma \in [0, 1]$, $\nu \in [0, 1)$. Then for $p$ sufficiently large, depending only on $B, d$, for $p' > p + 2$ and $\mu_0 \in S^{p'}$, there exists a unique energy-conserving solution $(\mu_t)_{t \geq 0}$ to the Boltzmann equation starting at $\mu_0$, which we write as $\mu_t = \phi_t(\mu_0)$. Moreover, for some constant $C = C(B, p, d)$,

(i) Whenever $\mu_0, \nu_0 \in S^{p'}$ satisfy the moment bound $\Lambda_{p+\gamma}(\mu_0, \nu_0) \leq a$, we have the continuity estimate

$$T_p(\phi_t(\mu_0), \phi_t(\nu_0)) \leq e^{Ca(1+t)} T_p(\mu_0, \nu_0).$$  \hspace{1cm} (15)

(ii) Whenever $\mu_0 \in S^l$, $l = p + 2 + \gamma$, the solution $\phi_t(\mu)$ is the $T_p$-limit of the solutions $\phi^K_t(\mu_0)$ to the K-cutoff Boltzmann Equations (BEK) starting at $\mu_0$, as the cutoff parameter $K \to \infty$. More precisely, if $\Lambda_{p+\gamma}(\mu_0) \leq a_1$, $\Lambda_l(\mu_0) \leq a_2$, for some $a_1, a_2 \geq 1$, then for all $K \geq K_0 = K_0(B, d)$, we have

$$T_p(\phi^K_t(\mu_0), \phi_t(\mu_0)) \leq e^{Ca_1(1+t)} a_2 t K^{1-1/\nu}. \hspace{1cm} (16)$$

We note that this is a much stronger wellposedness estimate than exists in the literature; see the discussion in the literature review below.

Our second result is to study the convergence of the noncutoff Kac process in the large number limit $N \to \infty$.

**Theorem 2.** (Propagation of Chaos) Let $B$ be as in Theorem 1. For all $N$, the $N$-particle Kac process defined by the generator (5) has uniqueness in law. Moreover, if $p, l$ are as in Theorem 1 and $q = 2p + 2\gamma + 4$, $a \geq 1$ and $t_{\text{fin}} \geq 0$, then whenever $\mu_0 \in S^q$ has a moment $\Lambda_q(\mu_0) \leq a$ and $\mu^N_t$ is a $N$-particle Kac process with initial data satisfying $\Lambda_q(\mu^N_0) \leq a$ almost surely, we have the estimate

$$E \left[ \sup_{t \leq t_{\text{fin}}} T_p(\phi_t(\mu_0), \mu^N_t) \right] \leq e^{Ca(1+t_{\text{fin}})} (\log N)^{1-1/\nu} + E \left[ T_p(\mu^N_0, \mu_0) \right]$$

\hspace{1cm} (17)

for some $C = C(B, p, q, d)$.

This estimate may also be understood as proving *propagation of chaos*; this will be discussed in the literature review below.
1.2. Summary of the key argument

Let us now discuss the key argument of this paper, which underlies both of the main results. Both Theorems 1, 2 will follow from a coupling at the level of the Kac process, which is presented in Lemma 8 and Corollary 1. The key point here is that the difference between the empirical measures $\mu^N_t$ of the noncutoff Kac process and $\mu^{N,K}_t$ of a suitably constructed cutoff Kac process can be controlled in $T_p$, uniformly in $N$, with the following errors:

(i) a multiple of $T_p(\mu^N_0,\mu^N_0)$, growing exponentially in time, and depending on the $(p+\gamma)$th moments of both processes;

(ii) an error depending on $K$, involving $l$th = $(p+2+\gamma)$th moments and vanishing as $K \to \infty$;

(iii) an error depending on $N$, depending on the $q$th = $(2p+4+2\gamma)$th moments, and vanishing as $N \to \infty$.

We first set up the coupling, following [14,20]. First, we reformulate the possible jumps in terms of parameters $z \in (0, \infty)$, $\varphi \in \mathbb{S}^{d-2}$, so that the binary collisions have the form $v \mapsto v + a(v, v_*, z, \varphi)$, $v_* \mapsto v_* - a(v, v_*, z, \varphi)$, for a suitable function $a : \mathbb{R}^d \times \mathbb{R}^d \times (0, \infty) \times \mathbb{S}^{d-2} \to \mathbb{R}^d$. This parametrisation is chosen so that the dependence of the rate on the kinetic factor $|v - v_*|^{\gamma}$ is absorbed into the dependence on $a$ through $z$, and the rate of collisions is given by $dz d\varphi dt$, which is constant in $z, \varphi, t$ and does not depend on the relative velocities. This will provide a natural coupling of two Kac processes with velocities $(V^i_1, \ldots, V^i_N)$, $(\tilde{V}^i_1, \ldots, \tilde{V}^i_N)$, by arranging that a collision between particles $i, j$ in one corresponds to a collision in the other, with the same value of $z$, if $z \leq K$, and where jumps in the coupled system $(\tilde{V}^i_1, \ldots, \tilde{V}^i_N)$ are suppressed if $z \geq K$, so that the coupled system is a cutoff Kac process. We must also use an ‘accurate Tanaka trick’ (Lemma 4), following [40], introducing a velocity dependent isometry $R$ which acts on the parameter $\varphi$, at each collision to compensate for the lack of continuity of the map $(v, v_*) \mapsto a(v, v_*, z, \varphi)$.

The problem with this coupling is that it does not, on its own, lead to a Grönwall estimate for the distance between the particle systems. Indeed, the rate of growth of the error is governed by

$$\frac{d}{dt} \bigg|_{t=0} \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N d_0^2(V^i_1, \tilde{V}^i_1) \right] = \frac{1}{N^2} \mathbb{E} \left[ \sum_{i=1}^N \mathcal{E}_0,K(V^i_1, V^i_0, \tilde{V}^i_1, \tilde{V}^i_0) \right]$$

where we recall that $d_0$ is the unweighted Euclidean distance $d_0(v, w) = |v - w|$, and $\mathcal{E}_0$ is given by

$$\mathcal{E}_0,K(v, v_*, \tilde{v}, \tilde{v}_*) := \int_{(0, \infty) \times \mathbb{S}^{d-2}} \left( d_0^2(v + a, \tilde{v} + \tilde{a}_K) - d_0^2(v, \tilde{v}) \right) dz d\varphi \quad (18)$$

with the shorthand $a = a(v, v_*, z, \varphi)$ for the jump of the original system and $\tilde{a} = a(v, v_*, \tilde{v}, R(v - v_*, \tilde{v} - \tilde{v})*)$ for the jump in the coupled system,
recalling that $R$ is the velocity-dependent isometry produced by the accurate Tanaka trick. However, attempting to estimate $E_0$ leads to terms of the forms

$$(1 + |v|^\gamma + |\tilde{v}|^\gamma)d^2_0(v, \tilde{v}) = d^2_p(v, \tilde{v}), \quad (1 + |v_*|^\gamma + |\tilde{v}_*|^\gamma)d^2_0(v, \tilde{v}_*).$$

Let us refer to, for example, [20, Lemmas 3.1, 3.3] for a similar calculation. Terms of the latter form can, after an integration in $v_*, \tilde{v}_*$ and using some moment estimates, be controlled by a multiple of $d^2_0(v, \tilde{v})$ and are thus amenable to a Grönwall estimate, but one must compensate for terms of the first kind, which prevent such an estimate; for example, [20] uses exponential moments of order $\alpha > \gamma$. This is where our weighted distance becomes important; repeating the same calculations with $d_p$ in place of $d_0$, we can break up the integrand of

$$E_{p,K}(v, v_*, \tilde{v}, \tilde{v}_*) := \int_{(0, \infty) \times S^{d-2}} \left( d^2_p(v + a, \tilde{v} + \tilde{a}_K) - d^2_p(v, \tilde{v}) \right) dz d\varphi \quad (19)$$

as

$$d^2_p(v + a, \tilde{v} + \tilde{a}_K) - d^2_p(v, \tilde{v}) = d^2_p(v, \tilde{v})(|v + a|^p - |v|^p + |\tilde{v} + \tilde{a}_K|^p - |\tilde{v}|^p) + (1 + |v + a|^p + |\tilde{v} + \tilde{a}_K|^p) \times (d^2_p(v + a, \tilde{v} + \tilde{a}_K) - d^2_p(v, \tilde{v})). \quad (20)$$

In the first term, the distance $d^2_0(v, \tilde{v})$ is constant, so the integral reduces to the change of the moment prefactors. This is, up to a reparametrisation, exactly the context of the evolution of polynomial moments and the Povzner inequalities, where it is well-known that a negative term appears [7,11,35], contributing a term like $- (|v|^{p+\gamma} + |\tilde{v}|^{p+\gamma})$ so that the contribution to $E_{p,K}$ is of the form

$$\lambda_p(1 + |v|^{p+\gamma} + |\tilde{v}|^{p+\gamma})d^2_0(v, \tilde{v}) = -\lambda_p d^2_{p+\gamma}(v, \tilde{v}) \quad (21)$$

for some $\lambda_p$ depending on $p$. The second line of (20) resembles the integral defining $E_{0,K}$, but is complicated by the $(z, \varphi)$-dependent moment prefactor; following the same arguments as for $E_{0,K}$, one finds the same kinds of terms $(1 + |v|^\gamma + |\tilde{v}|^\gamma)d^2_p(v, \tilde{v}) \leq C_d^2_{p+\gamma}(v, \tilde{v})$ and $(1 + |v_*|^\gamma + |\tilde{v}_*|^\gamma)d^2_p(v, \tilde{v}_*)$. The key calculation, which is summarised in Lemma 5 and carried out in Section 8, exactly shows that, provided $p$ is large enough, all of the terms of the first kind can be absorbed by the negative ‘Povzner term’ coming from the first line of (20), see also Remark 1. Regarding the second type of term, we show in Lemma 8 that the contributions can be controlled by a Grönwall-style estimate. All of this, keeping track of the errors due to $K < \infty, N < \infty$, leads to an estimate on the many-particle system (Lemma 8, Corollary 1) with the three kinds of error terms i-iii). described above.

From the Coupling to the Propagation of Chaos. Our strategy regarding Theorem 2 is that a coupling estimate, with errors only of the three forms described above, allows us to reduce the problem of convergence of the noncutoff Kac process to convergence of the cutoff Kac processes, which has already been the object of much study in the literature. We state some results on this convergence, which are essentially due to Norris [34], in Section 4, tracking the dependence in the cutoff parameter $K < \infty$ and ensuring that the number of moments is independent of $K$. 
Together with the results on the coupling developed in Lemma 8 and Corollary 1, these rapidly imply Theorem 2.

*From the Coupling to the Boltzmann Equation.* Let us remark that this coupling argument is not constrained to the Kac process, and can be applied to the Boltzmann equation directly, either at the level of Tanaka’s interpretation of the Boltzmann equation as a jump stochastic differential equation [40] or at the level of the partial differential equation, to derive uniqueness and stability directly at the level of the limit equation without involving the particle system. The argument for Tanaka’s stochastic interpretation was carried out in the author’s doctoral thesis after the first version of this work appeared [25, Chapter 4], and the argument at the level of the partial differential equation can be found in the closely-related context of the Landau equation in the joint work [22] by the author with Fournier, inspired by the present work. In either case, one needs an approximation argument to build a coupling of the Boltzmann (respectively Landau) equation; since this current work is also interested in the Kac process, it is natural to develop the coupling in this context and to use the Kac process as the approximation scheme. In Section 6, we will use the coupling of the Kac process to study the Boltzmann equation, building the coupling in the context of cutoff Kac processes as described above \( N < \infty, K < \infty \). In this case, the construction of the coupling is elementary, and we must take the limits to obtain the claimed stability of the noncutoff Boltzmann equation. We first use convergence results for Kac’s process from the literature [33,34] \( N \to \infty, K < \infty \) fixed, which are recalled in Section 4, to make errors of type iii) vanish, and then take \( K \to \infty \).

1.3. Plan of the paper

Our programme will be as follows:

(i) The remainder of this section is a literature review, discussing the existing results and techniques and the relationships to the current work.

(ii) In Section 2, we introduce an alternative parametrisation of possible jumps, and hence formulate the labelled Kac process discussed above as the solution to a stochastic differential equation driven by Poisson random measures. We state results on the wellposedness of the resulting stochastic differential equation and its relationship to the unlabelled Kac process.

(iii) Section 3 collects some results from the literature concerning the moments of the Kac process and the Boltzmann equation. We also prove a novel ‘concentration of moments’ result (Lemma 1) which shows that the moments of a Kac process remain in a fixed interval with probability converging to 1 in the large number limit \( N \to \infty \).

(iv) Section 4 recalls some results regarding the convergence of the Kac process and stability of the Boltzmann equation for the cutoff case \( K < \infty \), which we use as an intermediate step in later proofs. Since the proofs are not related to our coupling method, which is the main point of the present work, we do not give full-details, but in the interest of completeness an outline of the proofs can be found in Appendix A. The statements of intermediate steps, which
correspond closely to the original treatment by Norris [34], are given, as well as those aspects of the proof which significantly diverge from this treatment. A presentation of the full calculations can be found in the author’s doctoral thesis [25, Sections 3.2.1–3.1.2, pp. 112–128].

(v) Section 5 introduces the key coupling of cutoff and noncutoff Kac processes, making rigorous the informal description of the coupling above. We state our key estimate on the quantities $E_{p,K}$ in Lemma 5, and demonstrate the $N$-uniform coupling of the Kac processes in Lemma 8 and Corollary 1. The proof of the lemma is deferred until Section 8 for ease of readability.

(vi) In Section 6, we carefully transfer the previous coupling to the Boltzmann equation, use the convergence of the cutoff Kac process proven in Section 4. This proves Theorem 1.

(vii) In Section 7, we combine the previous results to deduce Theorem 2.

(viii) Section 8 presents the main calculations on our Tanaka coupling, deferred from Lemma 5.

(ix) Appendix A sketches the main points of the proof of the lemmata of Section 4, which are modifications of the results of Norris [34]. Appendix B deals with some technical issues concerning the well-posedness for labelled and unlabelled Kac processes, and the relationships between these.

1.4. Literature review

We will now briefly discuss related works and their relationship to our work.

1. Tanaka’s Coupling The key idea in our analysis is a coupling pioneered by Tanaka [40] in the case of cutoff Maxwell molecules ($\gamma = 0, \nu < 0$), who interpreted the Boltzmann equation in terms of a stochastic differential equation describing a ‘typical’ particle; this was generalised by Fournier and Méléard [13] to include the cases of non-Maxwellian molecules and used to show uniqueness for the Boltzmann equation with Maxwell molecules [41]. Since then, this coupling and analytic counterparts have been a popular idea in the treatment of the Boltzmann equation [4,16,17,20,36,42]. The ideas of the current paper have also been adapted to study the Boltzmann equation via this probabilistic representation in the author’s doctoral thesis [25, Chapter 4], and to study the the closely related Landau equation by Fournier and the author [22], [25, Chapter 5].

Let us mention two particular works to which our approach can be compared. The main calculations in Sections 5, 8 were inspired by Fournier and Mischler [20] on the Nanbu particle system, in which only one particle jumps at a time. In our notation, the cited paper produces estimates in $T_0$, which is the usual Wasserstein2 distance; the major novelty of this work is that, by working in $T_p$ for $p$ large enough, and working with the symmetric Kac process, we are able to obtain a desirable cancellation of ‘bad’ terms. Let us also remark that that the main result of Rousset [37] is very similar to the conclusion of Lemma 8 in obtaining a coupling of Kac processes with error uniform in $N$ in the case of Maxwell molecules, although the proofs are very different, and the ends to
which this is used are orthogonal to our programme. Rouset [37] is restricted to the case of Maxwell molecules, working in the metric $T_0 = w_2$, and the key argument in this case is to exploit a different negative term, strictly weaker than $\mathbb{E}[T_0(\mu_j^N, v_i^N)]$, to prove $N$-uniform convergence to equilibrium. Since the negative term exploited here is weaker even than the ‘good’ positive terms discussed above, it does not presently appear possible to combine this technique with our method, and we are restricted to the Boltzmann equation on fixed time intervals (as discussed below) by taking the limits $K \to \infty$, $N \to \infty$.

2. Wellposedness of the Boltzmann Equation The results expressed in Theorem 1 add to a long list of results concerning the wellposedness and stability of the Boltzmann equation. In the case of Maxwell molecules, including the noncutoff case, we refer the reader to [41,42]; in the case of hard spheres ($\gamma = 1, \nu < 0$), let us mention the works [3,12,29,30,32]. Most recently, Mischler and Mouhot [33] prove very strong ‘twice-differentiability’ in the initial data and exponential stability of the Boltzmann equation in the hard-spheres case measured in total variation distance, and the author obtained a uniform-in-time Wasserstein stability result in a previous work [24].

For the case of noncutoff hard potentials, the theory is substantially less complete. Fournier [15] examined the case where $|v - v_*|^{\gamma}$ is replaced by a bounded function $\Phi$, and results for the case of full hard potentials have been found by Desvillettes and Mouhot [10] and in the case of measure solutions by Fournier and Mouhot [17]. Let us note that the uniqueness and stability statement in Theorem 1 assumes only a finite number of moments, rather than a finite exponential moment $\langle e^{\epsilon |v|^{\gamma}}, \mu_0 \rangle < \infty$ as does the result of [17], which is recalled in Proposition 6 below; correspondingly, the quantitative stability result is stronger. The result of [10] requires the initial data $\mu_0$ to have a density $d\mu_0 / dv \in W^{1,1}(\mathbb{R}^d, dv)$, and so requires fewer moments than our results but much more regularity.

3. Propagation of Chaos for the Kac Process The sense in which Kac first proposed to relate his stochastic process to Boltzmann’s equation is through the propagation of chaos: he proposed that, if $V_t^N$ is a labelled Kac process, with symmetric initial conditions and $\mu_0 \in S$ such that marginal distribution of $(V^1_0, \ldots, V^K_0)$ is approximately $\mu_0^\otimes k$, then this approximation is propagated through time; the law of $(V^1_t, \ldots, V^K_t)$ is approximately $\phi_t(\mu_0)^\otimes k$, for any fixed $k, t$ in the regime where $N$ is large, and where the approximation is understood as the weak topology of measures on $(\mathbb{R}^d)^k$. This chaoticity property is equivalent to the convergence of the empirical measures [39]; quantitatively, the same arguments as in [24,33] show how the conclusion of Theorem 2 can be viewed as a quantitative estimate of this approximation. We now mention some existing works in this direction:

(i) For Maxwell molecules, results in this direction were obtained by McKean [31], Graham and Méléard [23] and Desvillettes, Graham and Méléard [8]. Strong results were obtained by Mischler and Mouhot [33], and close to optimal results were found by Cortez and Fontbona [5], in both cases uniformly in time.

(ii) Regarding the case of hard spheres, Mischler and Mouhot [33] obtained results decaying as $(\log N)^{-r}$ for some $r > 0$. Norris [34] obtained results
with optimal $N$ dependence, replacing the right-hand side in Theorem 2 with the optimal $N$ dependence $N^{-1/d}$ but which are not uniform in time, and the author [24] obtained results with close-to-optimal $N$ dependence and which are uniform in time.

(iii) For the case of non-cutoff hard potentials, we are not aware of any results on the Kac process, but mention some works on related models. Fournier and Guillin [18] consider a related particle system which approximates the Landau equation for hard potentials, and Fournier and Hauray [19] deal with this model for soft potentials. The work [20] which we have already mentioned considers the asymmetric Nanbu process in which only one particle jumps at a time, and shows propagation of chaos for this system for Maxwell molecules and hard potentials; a recent work of Salem [38] extends this to the case of soft potentials with a moderate singular angular singularity.

Our result on the propagation of chaos in Theorem 2 above comes from combining the key coupling described above, based on the calculations of [20], with the convergence as $N \to \infty$ with $K$ fixed which we use as an auxiliary result. The result we use in Lemma 2 is essentially due to Norris [34], but is marginally stronger than the corresponding result, replacing results with ‘high probability’ with expectation estimates thanks to the concentration of moments result Lemma 1; the key points of the proof are sketched in Appendix A, and the full calculations can be found in the author’s doctoral thesis [25, Sections 3.2.1–3.2.2, pp. 113–128]. We also remark that the rate obtained in Theorem 2 is equivalent to that of [8], and is likely very far from optimal; it may be possible to improve on this by using the regularising effect of grazing collisions [1,2,9] to improve the estimates in Section 4, but we will not explore this here.

2. A Jump Stochastic Differential Equation Associated to the Kac Process

As discussed in Section 1.2, we first introduce a parametrisation of collisions which will be helpful for our coupling argument, and correspondingly formulate a stochastic differential equation in $(\mathbb{R}^d)^N$ driven by Poissonian jump measures, which will correspond to the labelled Kac process discussed above. Introduce first measurable maps $\iota = (\iota_1, \ldots, \iota_{d-1}) : \mathbb{R}^d \to (\mathbb{R}^d)^{d-1}$ such that, for all $v \neq 0$, the set

$$\left\{ \frac{v}{|v|}, \frac{\iota_1(v)}{|v|}, \ldots, \frac{\iota_{d-1}(v)}{|v|} \right\}$$

is an orthonormal basis of $\mathbb{R}^d$, and $\iota(-v) = -\iota(v)$. With this choice of $\iota$, define $\Gamma : \mathbb{R}^d \times \mathbb{S}^{d-2} \to \mathbb{R}^d$ by

$$\Gamma(v, \varphi) = \sum_{j=1}^{d-1} \varphi_j \iota_j(v).$$

(23)
Let us also define
\[ H(\theta) = \int_0^{\pi/2} b(\cos x)dx, \quad \theta \in \left(0, \frac{\pi}{2}\right). \] (24)

Thanks to (3), \( H \) is now a bijection from \((0, \pi/2)\) to the ray \((0, \infty)\); let us write \( G \) for its inverse. We finally define, for distinct \( v, v_\ast \in \mathbb{R}^d \) and \( \varphi \in S^{d-2}, z > 0 \)
\[
\theta(v, v_\ast, z) = G \left( \frac{z}{|v - v_\ast|^2} \right); \tag{25}
\]
\[
a(v, v_\ast, z, \varphi) = -\frac{1 - \cos(\theta(v, v_\ast, z))}{2} (v - v_\ast) + \frac{\sin(\theta(v, v_\ast, z))}{2} \Gamma(v - v_\ast, \varphi). \tag{26}
\]

In the case \( v = v_\ast \), we set \( a(v, v_\ast, z, \varphi) = 0 \); we note that, by construction, \( a \) is antisymmetric in \( v, v_\ast \). Some estimates for the function \( G \) are established in Section 8.1.

With this parametrisation, we define a labelled Kac process to be the solution to an SDE with Poisson noise. For unordered pairs \( \{ij\} = \{ji\} \) of distinct indices \( i, j = 1, \ldots, N \), let \( \mathcal{N}^{ij} \) be independent Poisson random measures on \((0, \infty) \times S^{d-2} \times (0, \infty)\), with intensity \( 2N^{-1}dsd\varphi dz \). A labelled Kac process is then exactly a solution \( \mathcal{V}^N_t = (V^1_t, \ldots, V^N_t) \) to the system of stochastic differential equations
\[
V^i_t = V^i_0 + \sum_{j \neq i} \int_{(0,t] \times S^{d-2} \times (0,\infty)} a(V^i_{s_-}, V^j_{s_-}, z, \varphi) \mathcal{N}^{ij}(ds, d\varphi, dz), \tag{LK}
\]
where the index \( i \) runs over \( 1, \ldots, N \). The factor of 2 in the rate corresponds to working with unlabelled, rather than labelled, pairs of particles. Moreover, thanks to the antisymmetry of \( a \) in the first two arguments, and recalling that \( \mathcal{N}^{ij} = \mathcal{N}^{ji} \), we see that a jump in the \( i \)th particle \( V^i_t \neq V^i_{t-} \) matches a jump in some \( j \)th particle, \( j \neq i \).

Classically [28], weak solutions to the stochastic differential equation (\textbf{LK}) are Markov processes \( \mathcal{V}^N_t \) with the generator
\[
(G^L \widehat{F})(\mathcal{V}^N) = \frac{2}{N} \sum_{\{ij\}} \int_0^\infty dz \int_{S^{d-2}} d\varphi \left( \widehat{F} \left( \mathcal{V}^N + a(V^i, V^j, z, \varphi)(e_i - e_j) \right) - \widehat{F}(\mathcal{V}^N) \right)
\]
where the sum is over unordered pairs \( \{ij\} = \{ji\} \), and we use the notation, for \( 1 \leq i \leq N \) and \( h \in \mathbb{R}^d \), \( he_i \) is the vector \((0, \ldots, h, \ldots, 0)\) in \((\mathbb{R}^d)^N\) with \( h \) in the \( i \)th place. Thanks to the construction of \( G \) and \( a \), it is straightforward to check that the integral can be rewritten as
\[
(G^L \widehat{F})(\mathcal{V}^N) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \int_{S^{d-1}} \left( \widehat{F}(\mathcal{V}^N_{i,j,\sigma}) - \widehat{F}(\mathcal{V}^N) \right) d\sigma, \tag{28}
\]
where $V_{i,j}^N$ and $V_{i,j}^N$ denote the vector in $(\mathbb{R}^d)^N$ where the $i$th and $j$th coordinates have been updated according to (1); we note the strong similarity with (5).

In this way, we can think of (LK) as corresponding to the Kac process where each particle is assigned a label, and we call (weak) solutions to (LK) a labelled Kac process. We formalise the connection, and justify moving between the labelled and unlabelled dynamics, with the following proposition:

**Proposition 1.** (i) Suppose $V_i^N$ is a solution to the stochastic differential equation (LK). Then the empirical measures $\mu_t^N = \theta_N(V_t^N)$ are unlabelled Kac process.

(ii) Every Kac process arises in this way: if $(\tilde{\mu}_t^N, t \geq 0)$ is a Kac process starting at $\mu^N_0$, pick $V_0^N \in \theta_N^{-1}(\mu^N_0)$ uniformly at random. Then there exists a weak solution to the stochastic differential equation (LK), starting at $V_0^N$, such that $(\mu_t^N, t \geq 0) = (\theta_N(V_t^N), t \geq 0)$ has the same law as $(\tilde{\mu}_t^N, t \geq 0)$.

Since these technical points are rather tangential to the new ideas of this paper, the proof is deferred to Appendix B.

**Proposition 2.** For all $V_0^N \in S_N$, there exists a labelled Kac process $V_t^N, t \geq 0$, that is, a weak solution to (LK), starting at $V_0^N$. Therefore, if $\mu_0^N \in S_N$, then there exists an $N$-particle (unlabelled) Kac process starting at $\mu_0^N$.

**Proposition 3.** For all $V_0^N \in S_N$, the labelled Kac process $V_t^N, t \geq 0$ starting at $V_0^N$ has uniqueness in law. At the level of unlabelled dynamics if $\mu_0^N \in S_N$, then there exists a unique, in law, $N$-particle (unlabelled) Kac process starting at $\mu_0^N$.

Both proofs are again deferred until Appendix B. We emphasise that no results of this paper are necessary for the proofs, and so there is no circularity.

We also construct a cutoff version $V_t^{N,K} = (V_1^{1,K}, \ldots, V_t^{N,K})$ of these processes as follows. In analogy to the definition above, set

$$a_K(v, v^*, z, \varphi) = a(v, v^*, z, \varphi)1(\gamma \leq K|v - v^*|).$$

The $K$-cutoff version of (LK), corresponding to the cutoff kernel $B_K$ defined in (6) is now

$$V_t^{i,K} = V_i^{0,K} + \sum_{j \neq i} \int_{[0,t] \times \mathbb{R}^{d-2} \times (0,\infty)} a_K(V_s^{i,K}, V_s^{j,K}, z, \varphi) \mathcal{N}^{(ij)}(ds, d\varphi, dz).$$

(cLK)

In the notation above, $\theta_0(K) = H(K) \to 0$ as $K \to \infty$. Let us remark that the statements equivalent to Propositions 1, 2, 3 for the cutoff differential equation (cLK) and the corresponding cutoff Kac process $\mu_t^{N,K}$ are elementary, as in both cases the overall jump rates are uniformly bounded.
3. Moment Estimates for the Kac Process and Boltzmann Equation

We now present some results concerning the moment evolution for both the Kac process and the Boltzmann equation. Sections 3.1 and 3.2 summarise some moment estimates from the literature for the Kac process and Boltzmann equation in the cutoff and noncutoff cases, respectively, uniformly in both the number of particles $N$ and cutoff parameter $K$. The arguments are well-known for the Boltzmann equation, and have been proven for measure-valued solutions of the Boltzmann equation by Lu and Mouhot [30], and for the Kac process by Mischler and Mouhot [33] and Norris [34]. We give some details, in particular recalling the derivation of a Povzner inequality (36) and the corresponding appearance of a negative term (37), since this same phenomenon lies at the heart of the ‘Tanaka-Povzner estimate’ in Lemma 5, see the discussion in Section 1.2. In Section 3.3, we prove a novel ‘concentration of moments’ result.

3.1. Moment inequalities in the cutoff case

We start with some moment inequalities for the cutoff process and associated limit equation; we recall the notation $\phi^K_t(\mu_0)$ for the unique energy-conserving solution to (BE$_K$) for initial data $\mu_0 \in S$. The only novelty here is some care to ensure that the estimates are uniform in the cutoff parameter $K$ as soon as $K$ is bounded away from 0.

Proposition 4. (Moment Inequalities for the Cutoff Kac Process and Boltzmann Equation) We have the following bounds for polynomial velocity moments:

(i) Let $(\mu_t^{N,K})_{t \geq 0}$ be a $K$-cutoff Kac process on $N \geq 2$ particles, $K \geq 1$, started from $\mu_0^{N,K}$, and fix $q \geq p \geq 4$. Then there exists a constant $C(p,q) < \infty$, which does not depend on $K$, such that, for all $t \geq 0$,

$$
\mathbb{E} \left[ \Lambda_q(\mu_t^{N,K}) \right] \leq C(1 + t^{(p-q)/\gamma}) \Lambda_p(\mu_0^{N,K}). \tag{30}
$$

(ii) In the notation of the previous point, there exists a constant $C = C(p)$, also independent of $K$, such that for all $t_{\text{fin}} \geq 0$,

$$
\mathbb{E} \left( \sup_{0 \leq t \leq t_{\text{fin}}} \Lambda_p(\mu_t^{N,K}) \right) \leq (1 + C(p)t_{\text{fin}}) \Lambda_p(\mu_0^{N,K}). \tag{31}
$$

(iii) Let $k \geq 2$. In the notation of point i), we have the almost sure relation

$$
\mathbb{P} \left( \Lambda_k(\mu_t^{N,K}) \leq 2^{k+1} \Lambda_k(\mu_{t-}^{N,K}) \quad \text{for all } t \geq 0 \right) = 1. \tag{32}
$$

(iv) Let $p, q$ be as above, and let, and $\mu_0 \in \cup_{k \geq 2} S_k$. Then there exists a constant $C = C(p, q)$, which does not depend on $K \geq 1$, such that the solution $\phi^K_t(\mu_0)$ to the cutoff Boltzmann Equation satisfies

$$
\Lambda_q(\phi^K_t(\mu_0)) \leq C(1 + t^{(p-q)/\gamma}) \Lambda_p(\mu_0). \tag{33}
$$
Proof. Let us sketch the arguments leading to point items i)-ii), which are similar to those of [34, Proposition 3.1] but require some modification. It is convenient to work with a labelled, cutoff Kac process \( \mathcal{N}^{(ij)}_t \), and whose empirical measures are \( \mu_t^{N,K} \). Let us fix \( v, v_* \), and write \( x = |v - v_*|^\gamma, a = a(v, v_*, z, \varphi) \). We start from the straightforward calculation

\[
|v'|^2 = \left( \frac{1 + \cos G(z/x^\gamma)}{2} \right) |v|^2 + \left( \frac{1 - \cos G(z/x^\gamma)}{2} \right) |v_*|^2
\]

(34)

To bound the term on the second line, recall that \( |\Gamma(v - v_*, \varphi)| = |v - v_*| \leq |v| + |v_*| \) and observe that, since \( \Gamma(v - v_*, \varphi) \) is orthogonal to \( v - v_* \), \( v \cdot \Gamma(v - v_*, \varphi) = v_* \cdot \Gamma(v - v_*, \varphi) \), and so

\[
|v \cdot \Gamma(v - v_*, \varphi) | \leq \min(|v|, |v_*|) |\Gamma(v - v_*, \varphi)| \leq 2|v||v_*|.
\]

(35)

For any \( p \geq 4 \), we now raise both sides of (34) to the \( (p/2) \)th power to find

\[
|v + a|^p \leq \left( \frac{1 + \cos G(z/x^\gamma)}{2} \right)^{p/2} |v|^p + \left( \frac{1 - \cos G(z/x^\gamma)}{2} \right)^{p/2} |v_*|^p
\]

+ \( C_p (|v|^{p-1}|v_*| + |v_*|^{p-1}|v|) \sin G(z/x^\gamma) \).  

(36)

From this, and a similar inequality for \( |v_* - a|^p \), we obtain

\[
|v + a|^p + |v_* - a|^p - |v|^p - |v_*|^p \\
\leq -\beta(p, G(z/x^\gamma)) \left( |v|^p + |v_*|^p \right)
\]

+ \( C_p (|v|^{p-1}|v_*| + |v_*|^{p-1}|v|) \sin G(z/x^\gamma) \)

(37)

where

\[
\beta(p, \theta) = \left( 1 - \left( \frac{1 + \cos \theta}{2} \right)^{p/2} - \left( \frac{\sin \theta}{2} \right)^{p/2} \right).
\]

(38)

Since we consider only \( p \geq 4 \), we see that \( \beta(p, \theta) > 0 \) for all \( \theta \in (0, \pi/2) \). We conclude that, for any \( v, v_* \),

\[
\int_{(0,\infty) \times S^{d-2}} (|v + a|^p + |v_* - a|^p - |v|^p - |v_*|^p) 1(z \leq K|v - v_*|^\gamma) dz d\varphi
\]

\[
\leq -|v - v_*|^\gamma \left( \int_0^K \beta(p, G(z)) dz \right) (|v|^p + |v_*|^p)
\]

+ \( C|v - v_*|^\gamma \left( \int_0^K \sin G(z) dz \right) (|v|^{p-1}|v_*| + |v_*|^{p-1}|v|). \)

(39)

The coefficient multiplying the first term is bounded bounded away from 0 for \( K \geq 1 \), and, in particular, we can bound the right-hand side above by replacing this coefficient by a \( K \)-independent constant. For the second term, \( \int_0^K \sin G(z) dz \) is bounded, uniformly in \( K \), since \( \int_0^\infty \sin G(z) dz < \infty \). With this modification,
the same arguments as in [34, Proposition 3.1] lead to the first point of i). with $\Lambda_p$ replaced by $\langle |v|^p, \mu_{i}^{N,K} \rangle$. The conclusion follows on noting that, for some $C = C(p)$ and all $\mu \in \mathcal{S}$,

$$C^{-1} \Lambda_p(\mu) \leq \langle |v|^p, \mu \rangle \leq C \Lambda_p(\mu). \quad (40)$$

For the second point, we return to (37) to bound the jumps of $\Lambda_p(\mu_{i}^{N,K})$ by

$$\frac{C_p}{N} (\langle |V_{i,K}^{i-1}|V_{i,K}^{j,K} | - |V_{i,K}^{j-1}|V_{i,K}^{i,K} | \rangle \sin G \left( \frac{z}{|V_{i,K}^{i-1} - V_{i,K}^{j,K}|} \right) \right) \quad (41)$$

at points of $\mathcal{N}^{(ij)}$, since the first term is always negative. We now consider the process $A_t$ whose jumps are exactly the right-hand side, so that $A_t$ is increasing and

$$\sup_{s \leq t} \langle |v|^p, \mu_s^{N,K} \rangle \leq \langle |v|^p, \mu_0^{N,K} \rangle + A_t. \quad (42)$$

We now estimate

$$\mathbb{E}[A_t] \leq C \mathbb{E} \left[ \int_0^t \int_{\mathbb{R}^d} (|v|^p - |v_i| + |v_i|^p - |v_i|) |v - v_i|^{\gamma} \mu_s^{N,K} (dv) \mu_s^{N,K} (dv) \right] \quad (43)$$

where the constant is independent of $K$, due to the inclusion of the factor $\sin G$ in the definition of $A_t$. Simplifying, we see that

$$\mathbb{E}[A_t] \leq C_p \mathbb{E} \left[ \int_0^t \langle |v|^{p+\gamma-1}, \mu_s^{N,K} \rangle \right] \leq C_p \mathbb{E} \int_0^t \Lambda_p(\mu_s^{N,K}) ds \quad (44)$$

and the conclusion now follows, using the previous point to bound $\mathbb{E}[A_t]$. Item (iii) can be straightforwardly checked following the same calculations as [24, Lemma 2.1].

Regarding the Boltzmann Equation, item (iv), is standard in the literature for the cutoff Boltzmann equation, going back as far as Povzner and Bobylev, and these results can be checked by following the proofs in the quoted papers [30,33,34]. The key point here is that the constants arising in the proofs are independent of the cutoff $K$, which follows a similar pattern to the analysis of (39) above.

\[ \square \]

### 3.2. Moment inequalities in the noncutoff case

We will also use similar results for the noncutoff processes. In this case, where the available statement on well-posedness of the Boltzmann equation, or uniqueness in law for the Kac process, are weaker than in the cutoff case, we are careful that all the results cited are \textit{a priori} bounds, which are valid for any solution to the Boltzmann equation and any unlabelled Kac process, respectively.

**Proposition 5.** (Moment Inequalities for the Noncutoff Kac Process and Boltzmann Equation) We have the following bounds for polynomial velocity moments:
(i) Let $\mu^N_t$ be a noncutoff Kac process on $N \geq 2$ particles, and let $q \geq p \geq 4$. Then there exists a constant $C(p, q) < \infty$, such that, for all $t \geq 0$,
\[ \mathbb{E} \left[ \Lambda_q(\mu^N_t) \right] \leq C(1 + t^{(p-q)/\gamma}) \Lambda_p(\mu^N_0). \] (45)

(ii) In the notation of the previous point, there exists a constant $C = C(p)$ such that for all $t_{\text{fin}} \geq 0$,
\[ \mathbb{E} \left( \sup_{0 \leq t \leq t_{\text{fin}}} \Lambda_p(\mu^N_t) \right) \leq (1 + C(p)t_{\text{fin}}) \Lambda_p(\mu^N_0). \] (46)

(iii) Let $\mu^N_t$ be as in item i), and let $k \geq 2$. Then the bound (32) holds with $\mu^N_t$ in place of $\mu^N_t$.

(iv) Let $(\mu_t)_{t \geq 0} \subset S$ be a solution to the noncutoff Boltzmann equation (BE), and $p > 2$. Then there exists a constant $C = C(p) < \infty$ such that
\[ \sup_{t \geq 0} \Lambda_p(\mu_t) \leq C \Lambda_p(\mu_0). \] (47)

(v) In the notation of the previous point, for all $t > 0$, there exists $\varepsilon = \varepsilon(t) > 0$ such that
\[ \left\langle e^{\varepsilon(t)|v|^\gamma}, \mu_t \right\rangle < \infty. \] (48)

Proof. Concerning the Kac process, items (i)–(ii) follow as in the previous proposition, and item (iii) follows the same argument as in Proposition 4 and [24], although one must now be careful that $\mu^N_t$ is a jump process of infinite activity. Regarding the Boltzmann equation, the polynomial moment estimates go back to Desvillettes [7] and Wennberg [43] in the context of solutions admitting a density; the paper [30] extended the arguments to measure-valued solutions, and proved exponential moment creation [30, Lemma 4.1]. Let us also mention the recent work [21] by Fournier on exponential moment creation for this case, which demonstrates exponential moment creation of order strictly larger than $\gamma$ for the noncutoff case we consider here.

3.3. Concentration of moments

We also prove a result concerning the concentration of moments for the Kac process. The results above show uniform bounds on the expectation of moments, but to attain a bound of the form $\mathbb{P}(\Lambda_p(\mu^N_t) \leq b_N) \to 1$, we would need to take some sequence $b_N \to \infty$. The following result allows us to obtain such bounds with a single $b_N = b$ independent of $N$:

Lemma 1. (Concentration of Moments) Fix $p \geq 2$, and let $q \geq 2p + \gamma$. Then there exist constants $C_1(p)$, $C_2(p)$ such that, whenever $\mu^N_t$ is a (cutoff or non-cutoff) Kac
process on \(N\) particles satisfying an initial moment bound \(\Lambda_q(\mu_0^N) \leq a_3, a \geq 1\), then for all \(t_{\text{fin}} \geq 0\) and \(\varepsilon > 0\), we have the bound

\[
P \left( \sup_{t \leq t_{\text{fin}}} \langle |v|^p, \mu_t^N \rangle \geq \max(\langle |v|^p, \mu_0^N \rangle, C) + \varepsilon \right) \leq C t_{\text{fin}} a_3 N^{-1} \varepsilon^{-2}. \tag{49} \]

Define, for \(b \geq 1\),

\[
T_{b}^N = \inf \left\{ t \geq 0 : \Lambda_p(\mu_t^N) > \frac{b}{2^q+1} \right\}. \tag{50} \]

As a consequence of the estimate above, there exists \(C = C(p)\) such that, if the initial data has the moment estimates \(\Lambda_p(\mu_0^N) \leq a_1, \Lambda_q(\mu_0^N) \leq a_3\), then

\[
P(T_{C_{a_1}}^N \leq t_{\text{fin}}) \leq C t_{\text{fin}} a_3 N^{-1}. \tag{51} \]

The first statement here is somewhat sharper, and may be of independent interest; however, for applications later in the paper, it is useful to apply the second form, which absorbs some constants, without further comment.

**Proof.** Thanks to Proposition 1, it is sufficient to consider the case where \(V_t^N = (V_i^j, \ldots, V_i^N)\) is a labelled Kac process, and \(\mu_t^N\) are the associated empirical measures. Let us define

\[
M_t^N := \langle |v|^p, \mu_t^N \rangle - \langle |v|^p, \mu_0^N \rangle - \int_0^t \langle |v|^p, Q(\mu_s^N) \rangle \text{d}s
\]

\[
= \frac{1}{N} \int_{(0,t) \times \mathbb{R}^{d-2} \times (0,\infty)} \sum_{1 \leq i < j \leq N} H_{ij}^p(s, \varphi, z) \overline{N}^[ij](d\varphi, dz) \tag{52} \]

where we write

\[
H_{ij}^p(t, \varphi, z) := |V_{i}^j - a(V_i^j, V_i^j, z, \varphi)|^p - |V_i^j|^p
\]

\[
+ |V_{i}^j - a(V_i^j, V_i^j, z, \varphi)|^p - |V_j^i|^p \tag{53} \]

and

\[
\overline{N}^[ij](d\varphi, dz) = N^[ij](d\varphi, dz) - \frac{2}{N} d\varphi dz. \tag{54} \]

From the results of [6], \(M^N\) is a finite variation martingale, and thanks to Povzner estimates in the spirit of (39), for some \(\beta = \beta(p) > 0\) and all \(\mu \in S^{p+\gamma}\),

\[
\langle |v|^p, Q(\mu) \rangle \leq -\beta \langle |v|^{p+\gamma}, \mu \rangle + \beta^{-1} \langle |v|^p, \mu \rangle
\]

\[
\leq -\beta \langle |v|^p, \mu \rangle^{1+\gamma/p} + \beta^{-1} \langle |v|^p, \mu \rangle. \tag{55} \]

Set \(C_1 = \beta^{-2p/\gamma}\), so that the right-hand side of (55) is nonpositive as soon as \(\langle |v|^p, \mu \rangle \geq C_1\). Define \(T\) to be the stopping time

\[
T = \inf \left\{ t \geq 0 : \langle |v|^p, \mu_t^N \rangle > \max(C_1, \langle |v|^p, \mu_0^N \rangle) + \varepsilon \right\} \tag{56} \]
and on the event $T \leq t_{\text{fin}}$, define

$$T' = \sup \left\{ t < T : \langle |v|^p, \mu_t^N \rangle \leq \max(C_1, \langle |v|^p, \mu_0^N \rangle) \right\}. \quad (57)$$

This set is always nonempty, as it includes $0$, and we have

$$\limsup_{t' \uparrow T'} \langle |v|^p, \mu_t^N \rangle \leq \max(C_1, \langle |v|^p, \mu_0^N \rangle); \quad (58)$$

$$\langle |v|^p, \mu_t^N \rangle > \max(C_1, \langle |v|^p, \mu_0^N \rangle) \text{ for all } t \in (T', T]. \quad (59)$$

By the choice of $C_1$, it follows that

$$\int_{(T', T]} \langle |v|^p, Q(\mu_s^N) \rangle ds \leq 0 \quad (60)$$

and so, from (52), we must have $M_{T'}^N - M_{T^+}^N \geq \varepsilon$. Therefore, on the event $\{ T \leq t_{\text{fin}} \}$, we have the lower bound $\sup_{t \leq t_{\text{fin}}} |M_t^N| \geq \frac{\varepsilon}{2}$.

Let us now estimate $M^N$. From the analysis in [6], we have

$$\mathbb{E} \left[ |M_t^N|^2 \right] = \mathbb{E} \left[ \int_{(0, t] \times \mathbb{R}^{d-2} \times (0, \infty)} \frac{1}{N^2} \left( H_{ij}^N(s, \varphi, z) \right)^2 \frac{2}{N} ds d\varphi dz \right]. \quad (61)$$

To bound the integrand, we observe that $|a|^2 = |V_i^j - V_j^i|^2(1 - \cos G(z/x^\gamma))$, where $x = |V_i^j - V_j^i|$. We therefore obtain

$$\begin{align*}
( |V_i^j + a|^p - |V_i^j|^p )^2 \\
\leq C(p)(1 + |V_i^j + a|^{p-1} + |V_i^j|^p)|a|^2 \\
\leq C(p)(1 + |V_i^j|^{2p-2} + |V_j^i|^2|V_i^j|^{2p-2}) |V_i^j - V_j^i|^2(1 - \cos G(z/x^\gamma)) \\
\leq C(p)(1 + |V_i^j|^2 + |V_j^i|^2)G(z/x^\gamma)^2. \quad (62)
\end{align*}$$

Thanks to the estimates in Section 8.1, it follows that $\int_0^{\infty} G^2 dz < \infty$, and in particular, the integral on the right-hand side of (61) is finite. Using a similar computation for $(|V_i^j - a|^p - |V_i^j|^p)^2$, we obtain

$$\int_{\mathbb{R}^{d-2} \times (0, \infty)} \left( H_{ij}^N(s, \varphi, z) \right)^2 d\varphi dz \leq C(p)(1 + |V_i^j|^{2p} + |V_j^i|^{2p})x^\gamma \quad (63)$$

$$\leq C(p)(1 + |V_i^j|^{2p+\gamma} + |V_j^i|^{2p+\gamma}).$$

Returning to (61), we sum over pairs $i, j$ to obtain, for some $C_2 = C_2(p)$,

$$\mathbb{E} \left[ |M_t^N|^2 \right] \leq \frac{C_2}{16N} \mathbb{E} \left[ \int_0^t \Lambda_{2p+\gamma}(\mu_s^N) ds \right]. \quad (64)$$
By the choice of $\mu^N_0$ and moment propagation results above, the right-hand side is at most $C_2 t_{\text{fin}} a_3 / 16 N$. The first item now follows by using Doob’s $L^2$ inequality to bound $\mathbb{E}[\sup_{t \leq t_{\text{fin}}} |M^N_t|^2]$, and Chebychev’s inequality to bound the probability

$$\mathbb{P} \left( \sup_{t \leq t_{\text{fin}}} |M^N_t| \geq \frac{\varepsilon}{2} \right) \leq 16 \varepsilon^{-2} \mathbb{E} \left[ |M^N_{t_{\text{fin}}}|^2 \right].$$  (65)

The second item is a largely trivial reformulation of the first, noting that $\frac{\Lambda_1}{\langle v^p, \mu^N \rangle}$ is bounded, and since $a_1 \geq 1$ and $a_1 \geq \langle |v|^p, \mu^N \rangle$, we can choose $C = C(p)$ so that, on the event $\{ \Lambda_1(p, \mu^N) > C a_1 / 2 t_{\text{fin}}^{2 + \gamma} \}$, we also have $\langle |v|^p, \mu^N \rangle > \max(\langle |v|^p, \mu^0 \rangle, C_1) + 1$.  \(\square\)

4. Analysis of the Cutoff Kac Process and Boltzmann Equation

In this section, we will present some results concerning the cutoff Kac process and Boltzmann equation, which we will use as an intermediate result to go from our coupling argument at the level of the Kac process to the Boltzmann equation, and to deduce the convergence of the noncutoff Kac process. We use the results that follow, essentially due to Norris [34], as these techniques ensure that the number of moments required for our convergence result does not depend on the cutoff parameter $K$. As already mentioned, existence and uniqueness is well-established in the literature for the cutoff case; see, for instance, [30] and the references therein. The next result builds on [34, Theorem 1.1], and quantifies the rate of this convergence.

**Lemma 2.** (Convergence of the Cutoff Kac Process) Let $p \geq 0$ and $q > \max(4 + 3 \gamma, p + 2)$. Then there exists $C = C(p, q, d), \alpha = \alpha(d, p, q) > 0$, $\beta = \beta(d, p, q, \gamma) > 0$ such that, whenever $a \geq 1$, $\mu_0 \in S$ and $\mu^N,K_t$ is a $K$-cutoff Kac with $K \geq 1$ and initial moment estimates

$$\Lambda_q(\mu_0) \leq a, \quad \mathbb{P}(\Lambda_q(\mu^N,0,K) \leq a) = 1,$$  (66)

then we have the convergence estimate, for all $t_{\text{fin}} \geq 0$,

$$\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} T_p^2 \left( \mu^N,K_t, \phi^K_t(\mu_0) \right) \right] \leq \left( N^{-\alpha} + \mathbb{E} \left[ T_p \left( \mu^N,0,K, \mu_0 \right) \right]^\beta \right) e^{C a K (1 + t_{\text{fin}})}.$$

(67)

The following lemma, which may be of interest in its own right, arises as a step towards the above result:

**Lemma 3.** (Wasserstein Stability for Cutoff Boltzmann Equation) For all $p \geq 0, p' > p + 2$, there exists a constant $C = C(p, p') < \infty$ and $\beta = \beta(p, p', \gamma) > 0$ such that, for all $\mu, \nu \in S^{p'}$ and all $K < \infty$, we have the stability, for $t \geq 0$,

$$T_p^2 \left( \phi^K_t(\mu), \phi^K_t(\nu) \right) \leq \exp \left( C K (1 + t) \Lambda_{2 + \gamma}(\mu, \nu) \right) T_p(\mu, \nu)^\beta \Lambda^2_p(\mu, \nu).$$ (68)
Since the proofs of these lemmata rely on very different techniques than those of this paper, and are essentially already known, we omit the proofs from the main text. However, in the interest of keeping the paper self-contained, and since some points do diverge from the treatment of [34], the important points of the argument are reproduced in Appendix A, and details can be found in the author’s doctoral thesis [25, Sections 3.2.1–3.2.2, pp. 113–128]. The results cited here, up to working with Hölder-equivalent distances, Corollaries 3.13 and 3.14.

5. Tanaka Coupling of the Kac Processes

In this section, we will exhibit the key coupling of Kac processes, and a family of lemmata which control how fast the distance between the coupling can grow.

5.1. Accurate Tanaka’s Trick

We begin with the an ‘accurate Tanaka Lemma’, which generalises that of [16]. As discussed in in Section 1.2 in the introduction, this isometry will be necessary to build the coupling of the Kac processes, in order to compensate for the fact that the maps \( v, v_\ast \mapsto a(v, v_\ast, \varphi, z) \) constructed in Section 2 are not continuous. Our result is slightly more general than the cited result, in that we allow any \( d \geq 3 \), while the result cited applies for only \( d = 3 \).

**Lemma 4. (Accurate Tanaka’s Trick)** There exists a measurable function \( R : \mathbb{R}^d \times \mathbb{R}^d \to \text{Isom}(S^{d-2}) \) such that, for all \( X, Y \in \mathbb{R}^d \) and \( \varphi \in S^{d-2} \), we have

\[
\Gamma(X, \varphi) \cdot \Gamma(Y, R\varphi) = \varphi_1^2(X \cdot Y) + (1 - \varphi_1^2)|X||Y| \geq X \cdot Y. \tag{69}
\]

Here, \( \varphi_1 \) denotes the first coordinate of \( \varphi \in S^{d-2} \subset \mathbb{R}^{d-1} \).

**Proof.** First, the case where either \( X, Y = 0 \) is vacuous and can be omitted. Let us write, throughout, \( S_X \) for the set

\[
S_X = \{ u \in \mathbb{R}^d : |u| = |X|, u \cdot X = 0 \}. \tag{70}
\]

By considering separately the cases where \( X, Y \) are and are not colinear, we observe that we may choose \( j_X^1, j_Y^1 \) such that

\[
\dim \text{Span}(X, Y, j_X^1, j_Y^1) = 2, \quad j_X^1 \in S_X, j_Y^1 \in S_Y; \quad j_X^1 \cdot j_Y^1 = X \cdot Y. \tag{71}
\]

With some care, the map \((X, Y) \mapsto (j_X^1, j_Y^1)\) can further be constructed to be measurable. We now construct, in a measurable way, \( u_2, \ldots, u_{d-1} \) as an orthonormal basis for \( \text{Span}(X, j_X^1)^\perp = \text{Span}(Y, j_Y^1)^\perp \), and set

\[
j_X^2 = |X|u_2, \quad j_X^3 = |X|u_3, \ldots, \quad j_X^{d-1} = |X|u_{d-1}; \tag{72}
\]

\[
j_Y^2 = |Y|u_2, \quad j_Y^3 = |Y|u_3, \ldots, \quad j_Y^{d-1} = |Y|u_{d-1}. \tag{73}
\]
Now, \( \{j_X^1, \ldots, j_X^{d-1}\} \) are orthogonal, and lie in \( S_X \), so there is a unique isometry \( P_X \in \text{Isom}(\mathbb{S}^d_X) \) such that
\[
\Gamma(X, P_X e_k) = j_X^k, \quad k = 1, \ldots, d - 1
\] (74)
and similarly for \( P_Y \). We now observe, for all \( \varphi \in \mathbb{S}^{d-2} \),
\[
\Gamma(X, P_X \varphi) \cdot \Gamma(Y, P_Y \varphi) = \sum_{k=1}^{d-1} \phi_k^2 j_X \cdot j_Y = \phi_1^2 (X \cdot Y) + (1 - \varphi_1^2)|X||Y|
\geq \phi_1^2 (X \cdot Y) + (1 - \varphi_1^2)(X \cdot Y)
= X \cdot Y
\] (75)
which implies the result when we define \( R(X, Y) = P_Y P_X^{-1} \). \( \square \)

5.2. Tanaka–Povzner Lemmata

The key tool at the heart of our results is a variant of some calculations in [20, Lemmata 3.1, 3.3]. As described in Section 1.2 in the introduction, the key point is the appearance of a large negative term, similar to those arising in \((36, 39)\) from the Povzner inequalities in Section 3, which ensures the cancellation of ‘bad’ terms and leads to a Grönwall inequality.

**Lemma 5.** Let us write, for \( v, \tilde{v}, v_*, \tilde{v}_* \in \mathbb{R}^d, \ z \in (0, \infty), \varphi \in \mathbb{S}^{d-2} \) and \( K < \infty \),
\[
a = a(v, v_*, z, \varphi); \quad v' = v + a; \quad v'_K = v + a_K(v, v_*, z, \varphi);
\]
\[
\tilde{a}_K = a_K(\tilde{v}, \tilde{v}_*, z, R(v - v_*, \tilde{v} - \tilde{v}_*)\varphi); \quad \tilde{v}'_K = \tilde{v} + \tilde{a}_K
\] (76) (77)
for the isometry \( R(v - v_*, \tilde{v} - \tilde{v}_*) : \mathbb{S}^{d-2} \to \mathbb{S}^{d-2} \) constructed in Lemma 4. Recalling the definition \( d_p^2(v, w) := (1 + |v|^p + |w|^p)|v - w|^2 \), define
\[
E_{p,K}(v, \tilde{v}, v_*, \tilde{v}_*) = \int_0^\infty dz \int_{\mathbb{S}^{d-2}} d\varphi \left( d_p^2(v', \tilde{v}'_K) - d_p^2(v, \tilde{v}) \right).
\] (78)
Let us define, for \( p \geq 2 \),
\[
\lambda_p = \int_0^{\pi/2} \left( 1 - \left( \frac{1 + \cos \theta}{2} \right)^{p/2} \right) \beta(\theta) d\theta.
\] (79)
Then there exists \( K_0(p) \), constants \( c = c(G, d) \) and \( C = C(G, d, p) \), such that, whenever \( K \geq K_0(p) \), we have
\[
E_{p,K}(v, \tilde{v}, v_*, \tilde{v}_*) \leq \left( c + \left( c - \frac{\lambda_p}{2} \right) |v|^{p+\gamma} + \left( c - \frac{\lambda_p}{2} \right) |\tilde{v}|^{p+\gamma} \right) |v - \tilde{v}|^2
+ c \left( |v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma} \right) |v_* - \tilde{v}_*|^2
+ C \left( |v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma} \right) d_p^2(v, \tilde{v})
+ C \left( |v|^{p+\gamma} + |\tilde{v}|^{p+\gamma} \right) d_p^2(v_*, \tilde{v}_*)
+ C K^{1-l/\nu} (1 + |v|^l + |v_*|^l + |\tilde{v}|^l + |\tilde{v}_*|^l)
\] (80)
where \( l = p + 2 + \gamma \).
Remark 1. This lemma exactly corresponds to the sketch analysis in Section 1.2 in the introduction. Without the negative terms proportional to $-\frac{\lambda_p}{2}$ in the first line, the term on the first line cannot be controlled by $d_p(v,\tilde{v})$, and we would not find a Grönwall estimate. Our strategy is to exploit the negative ‘Povzner’ terms in the first line, corresponding to (21). The other terms in the four lines are exactly the two kinds of terms described in the paragraph under (21), and the final term is the error induced by coupling a cutoff to a non-cutoff system. We insist here that the terms of the form $|v|^{p+\gamma} + |\tilde{v}|^{p+\gamma}|v - \tilde{v}|^2$ are multiplied by constants $c = c(G, d, p)$ independent of $p$, so that they can be absorbed by the Povzner term if $\lambda_p$ is sufficiently large, which for the noncutoff case can be guaranteed by making $p$ sufficiently large.

We will also use the following variants, which will be used to prove a local uniform estimate on our coupling:

Lemma 6. In the notation of the previous lemma, define also

$$Q_K(v, \tilde{v}, v_*, \tilde{v}_*) = \int_0^\infty dz \int_{S^{d-2}} d\varphi \left( d_p^2(v', \tilde{v}') + d_p^2(v'_*, \tilde{v}_*') - d_p^2(v, \tilde{v}) - d_p^2(v_*, \tilde{v}_*) \right)^2. \quad (81)$$

Then, for some $C = C(G, d, p)$, we have

$$Q_K(v, \tilde{v}, v_*, \tilde{v}_*) \leq C(1 + |v|^{2l} + |v_*|^{2l} + |\tilde{v}|^{2l} + |\tilde{v}_*|^{2l}). \quad (82)$$

where $l = p + 2 + \gamma$ is as in Lemma 5.

5.3. Coupling of the Kac process

We now present our coupling of the Kac processes. Let $\mathcal{V}_t^N = (V_t^1, \ldots, V_t^N)$ be a noncutoff Kac process, and let $\mathcal{N}^{(ij)}$, $1 \leq i \neq j \leq N$, be the Poisson random measures of intensity $2N^{-1} d\varphi dz$ driving $\mathcal{V}_t^N$, so that

$$V_t^i = V_0^i + \sum_{j \neq i} \int_{(0,1] \times S^{d-2} \times (0, \infty)} a(V_t^j, V_t^j, z, \varphi) \mathcal{N}^{(ij)}(ds, d\varphi, dz),$$

$$i = 1, \ldots, N. \quad (83)$$

Let us fix $\mathcal{V}_t^N = (V^1_0, \ldots, V^N_0)$ and define $\mathcal{V}^i_N, K = (\tilde{V}^i_1, K, \ldots, \tilde{V}^i_N, K)$ by

$$\tilde{V}_t^i, K = \tilde{V}_0^i, K + \sum_{j \neq i} \int_{(0,1] \times S^{d-2} \times (0, \infty)} a_K(\tilde{V}_t^j, K, \tilde{V}_t^j, z, \tilde{V}^i, K, R_t^i, z) \mathcal{N}^{(ij)}(ds, d\varphi, dz);$$

$$R_t^i, j \coloneqq R(V_t^i - V_t^j, \tilde{V}_t^i, K - \tilde{V}_t^j, K), \quad (84)$$

where $R : \mathbb{R}^d \times \mathbb{R}^d \to \text{Isom}(S^{d-2})$ is the isometry constructed in Lemma 4. We remark first that the rates of $\mathcal{N}^{(ij)}$ are all finite on the support of $a_K$, so that the
stochastic differential equation (84, 85) is really a recurrence relation; in particular, \( \tilde{V}^{N,K}_t \) is uniquely defined by the above equations. Next, we claim that \( \tilde{V}^{N,K}_t \) is a \( K \)-cutoff Kac process on \( N \) particles; this is the content of the following lemma, which is adapted from a similar claim [20, Proposition 4.4]:

**Lemma 7.** Let \( V^N_t \) be a noncutoff Kac process, and fix \( \tilde{V}^{N,K}_0 \). Then the process \( (\tilde{V}^{N,K}_t)_{t \geq 0} \) constructed by (84) is a cutoff Kac process starting at \( \tilde{V}^{N,K}_0 \).

We will omit the proof, since it is identical to the above cited result. The key point is that the transformations \( R^{ij}_t \) are previsible, and preserve the uniform measure \( d\phi \) on \( S_{d-2} \). We now give our main result on the coupling.

**Lemma 8.** (Convergence of the Tanaka Coupling) There exists \( p_0 = p_0(G, d) \) and, for \( p > p_0 \), there exists \( K_0 = K_0(G, p, d) \) such that, whenever \( p > p_0 \) and \( K > K_0 \), we have the following estimates.

Let \( V^N_t \) be a noncutoff labelled Kac process and \( \tilde{V}^{N,K}_0 \in S_N \). Let \( \tilde{V}^{N,K}_t \) be the cutoff Kac process constructed in (84), and define

\[
\bar{d}^2_p(t) := \frac{1}{N} \sum_{i=1}^{N} d^2_p \left( V^i_t, \tilde{V}^{i,K}_t \right). \tag{86}
\]

Suppose the initial data \( V^N_0, \tilde{V}^{N,K}_0 \) are such that the associated empirical measures \( \mu^N_0, \tilde{\mu}^{N,K}_0 \) satisfy moment bounds

\[
\Lambda_l \left( \mu_0^N, \tilde{\mu}_0^{N,K} \right) \leq a_2; \tag{87}
\]
\[
\Lambda_q \left( \mu_0^N, \tilde{\mu}_0^{N,K} \right) \leq a_3; \tag{88}
\]

with \( l \) as in Lemma 5 and \( q = 2l \), and for some \( a_2, a_3 > 1 \). Fix \( b > 1 \), and let \( T^N_b \) be the stopping time (50) for the empirical measures \( \mu^N_t \) of \( V^N_t \), with \( p + \gamma \) in place of \( p \), and similarly \( T^{N,K}_b \) for \( V^{N,K}_t \). Then there exists \( C = C(p, G, d) \) such that, for all \( t \geq 0 \),

\[
\mathbb{E} \left[ \bar{d}^2_p(t) \right] \leq e^{Cb(1+t)} \left( \bar{d}^2_p(0) + a_2 t K^{1-1/v} \right) + a_3 C \mathbb{P} \left( T^N_b \land T^{N,K}_b \leq t \right)^{1/2} \tag{89}
\]

and, for all \( t_{\text{fin}} \geq 0 \),

\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} \bar{d}^2_p(t) \right] \leq e^{Cb(1+t_{\text{fin}})} \left( \bar{d}^2_p(0) + a_2 t K^{1-1/v} + \frac{Ca_3 t_{\text{fin}}^{1/2}}{N^{1/2}} \right) + a_3 C (1 + t_{\text{fin}}) \mathbb{P} \left( T^N_b \land T^{N,K}_b \leq t \right)^{1/2}. \tag{90}
\]

Before giving the proof of these estimates, let us demonstrate a corollary at the level of unlabelled cutoff processes which is natural for later analyses.
Corollary 1. Let \( p, q, l, K_0, C \) be as in Lemma 8, and let \( K' \geq K > K_0(G, p, d) \). Let \( \mu_0^N, \tilde{\mu}_0^N \) be random variables taking values in \( S_N \), with almost sure moment bounds

\[
\Lambda_{p+q} \left( \mu_0^N, \tilde{\mu}_0^N \right) \leq a_1; \quad \Lambda_1 \left( \mu_0^N, \tilde{\mu}_0^N \right) \leq a_2; \quad \Lambda_\theta \left( \mu_0^N, \tilde{\mu}_0^N \right) \leq a_3.
\]

The following hold.

(i) (Coupling Noncutoff and Cutoff) There exists a coupling of a noncutoff Kac process \( \mu_t^N \) starting at \( \mu_0^N \) and a \( K \)-cutoff Kac process \( \tilde{\mu}_t^{N,K} \) starting at \( \tilde{\mu}_0^N \) such that, for all \( t \geq 0 \),

\[
\mathbb{E} \left[ T_p \left( \mu_t^N, \tilde{\mu}_t^{N,K} \right) \right] \leq e^{C_1(1+t)} \left( T_p \left( \mu_0^N, \tilde{\mu}_0^{N,K} \right) + a_2 K^{1-1/v} \right) + C_3 K t^{1/2} \]

and, for all \( t_{\text{fin}} \geq 0 \),

\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \mu_t^N, \tilde{\mu}_t^{N,K} \right) \right] \leq e^{C_1(1+t_{\text{fin}})} \left( T_p \left( \mu_0^N, \tilde{\mu}_0^{N,K} \right) + a_2 t K^{1-1/v} \right) + C_3 \left( 1 + t_{\text{fin}} \right)^2 N^{1/2} \]

(ii) (Coupling with different Cutoffs) There exists a coupling of cutoff Kac processes \( \mu_t^{N,K}, \tilde{\mu}_t^{N,K'} \) with cutoff parameters \( K, K' \) respectively, starting at \( \mu_0^N = \mu_0^N, \tilde{\mu}_0^{N,K'} = \tilde{\mu}_0^N \), such that

\[
\mathbb{E} \left[ T_p \left( \mu_t^{N,K}, \tilde{\mu}_t^{N,K'} \right) \right] \leq e^{C_1(1+t)} \left( T_p \left( \mu_0^N, \tilde{\mu}_0^{N,K} \right) + a_2 t K^{1-1/v} \right) + a_3 C N^{-1/2} t. \]

Proof. Let \( p, K, \mu_0^N, \tilde{\mu}_0^N, a_1, a_2, a_3 \) be as given in the statement, and choose \( \gamma_0^N \in \theta_N^{-1}(\mu_0^N), \tilde{\gamma}_0^N \in \theta_N^{-1}(\tilde{\mu}_0^N) \), recalling that \( \theta_N \) is the map taking an \( N \)-tuple to its empirical measure, such that

\[
T_p(\mu_0^N, \tilde{\mu}_0^N) = \frac{1}{N} \sum_{i=1}^N d_p^2(\gamma_i^0, \tilde{\gamma}_i^0). \]

Now, let \( \gamma_i^N \) be a noncutoff labelled Kac process starting at \( \gamma_i^N \), and write \( \mu_t^N = \theta_N(\gamma_i^N) \) for the process of empirical measures. Let \( \tilde{\gamma}_i^{N,K} \) be the Kac processes constructed by Lemma 8 for the initial data \( \tilde{\gamma}_0^N \) with cutoff parameter \( K \), and let \( \tilde{\mu}_t^{N,K} \) be the associated empirical measures. We observe that

\[
T_p \left( \mu_t^N, \tilde{\mu}_t^{N,K} \right) \leq \frac{1}{N} \sum_{i=1}^N d_p^2 \left( \gamma_i^N, \tilde{\gamma}_i^{N,K} \right) = \tilde{d}_p^2(t).
\]
which we control by the previous lemma to obtain, for some $C$ and all $t \geq 0$, $b > 1$
\[ \mathbb{E}\left[T_p \left(\mu^N_i, \tilde{\mu}^N_{i,K}\right)\right] \leq e^{Cb(1+t)} \left(T_p \left(\mu^N_0, \tilde{\mu}^N_0\right) + a_2K^{1-1/\nu}\right) + Ca_3\mathbb{P}\left(T^N_b \land T^N_{b,K} \leq t\right)^{1/2}. \quad (97) \]

where $T^N_b, T^N_{b,K}$ are as above. Now, taking $b = Ca_1$ for some large $C = C(p)$, we use Lemma 1 to control the final term and obtain, for some $C$,
\[ \mathbb{P}\left(T^N_{Ca_1} \land T^N_{Ca_1} \leq t\right) \leq \mathbb{P}\left(T^N_{Ca_1} \leq t\right) + \mathbb{P}\left(T_{Ca_1} \leq t\right) \leq Cta_3N^{-1} \quad (98) \]
so that (97) becomes (92) as desired. We obtain (93) from (90) for the same processes $\mu^N_i, \tilde{\mu}^N_{i,K}$ in exactly the same way.

Regarding item (ii), the proof works in the same way. Let us pick $V^N_0 \in \theta^{-1}_N(\mu^N_0)$ and $\tilde{V}^N_0 \in \theta^{-1}_N(\tilde{\mu}^N_0)$, and construct a noncutoff labelled process $V^N_t$ starting at $V^N_0$ and a $K'$-cutoff $\tilde{V}^N_{t,K'}$ starting at $\tilde{V}^N_0$. We take $\mu^N_t, \tilde{\mu}^N_{t,K'}$ to be the associated empirical measures, which are (unlabelled) Kac processes. We now repeat this argument to construct a $K$-cutoff process $V^N_t, K^N_t$ starting at the same point $V^N_0 = V^N_{0,K}$, and let $\mu^N_t, \tilde{\mu}^N_{t,K}$ be the associated empirical measures. The same argument as in item i). establishes controls on
\[ \mathbb{E}\left[T_p \left(\mu^N_t, \tilde{\mu}^N_{t,K}\right)\right]; \quad \mathbb{E}\left[T_p \left(\mu^N_t, \mu^N_{t,K}\right)\right]. \quad (99) \]
Recalling the relaxed triangle inequality (12), we combine these to find the desired estimate. \qed

5.4. Proof of Lemma 8

We now give the proof of Lemma 8, which we deferred earlier.

Proof. Let $p \geq 0$ to be decided later, and consider the processes
\[ M^i_t = d^2_p(V^i_t, V^i_{t,K}) - d^2_p(V^i_0, \tilde{V}^i_0) - \frac{2}{N} \int_0^t \sum_{j=1}^N \mathcal{E}_{p,K}(V^i_s, V^i_{s,K}, V^j_s, V^j_{s,K})ds \quad (100) \]
for $1 \leq i \leq N$, and their average
\[ \overline{M}_t = \frac{1}{N} \sum_{i=1}^N M^i_t = \overline{d}^2_p(t) - \int_0^t \mathcal{E}_{p,K}(s)ds \quad (101) \]
where we define
\[ \mathcal{E}_{p,K}(t) := \frac{2}{N^2} \sum_{i,j=1}^N \mathcal{E}_{p,K}(V^i_t, \tilde{V}^i_{t,K}, V^j_t, \tilde{V}^j_{t,K}). \quad (102) \]
By classical results in the theory of Markov chains [6], each $M^i_t$ is a martingale, and hence so is $\overline{M}$. By Lemma 5, provided $K$ is large enough, depending on $G$, $p$, $d$, we have, for some $c = c(G, d)$, $C = C(G, d, p)$,

$$
\mathcal{E}_{p,K}(V^i_t, \overline{V}^{i,K}_t, V^j_t, \overline{V}^{j,K}_t) \leq \left( c + \left( c - \frac{\lambda_p}{2} \right) |V^i_t|^{p+\gamma} \\
+ \left( c - \frac{\lambda_p}{2} \right) |\overline{V}^{i,K}_t|^{p+\gamma} \right) |V^i_t - \overline{V}^{i,K}_t|^2 \\
+ c \left( |V^j_t|^{p+\gamma} + |\overline{V}^{j,K}_t|^{p+\gamma} \right) |V^j_t - \overline{V}^{j,K}_t|^2 \\
+ C \left( |V^i_t|^{p+\gamma} + |\overline{V}^{i,K}_t|^{p+\gamma} \right) d_p^2(V^i_t, \overline{V}^{i,K}_t) \\
+ C \left( |V^j_t|^{p+\gamma} + |\overline{V}^{j,K}_t|^{p+\gamma} \right) d_p^2(V^j_t, \overline{V}^{j,K}_t) \\
+ C K^{1-1/v} \left( 1 + |V^i_t|^l + |V^j_t|^l + |\overline{V}^{i,K}_t|^l + |\overline{V}^{j,K}_t|^l \right).
$$

Let us now take the average over all $i, j$. The first two lines can be absorbed together, as can the third and the fourth; for some new constants $c$, $C$ with the same dependence as above,

$$
\mathcal{E}_{p,K}(t) \leq \frac{1}{N} \sum_{i=1}^N \left( c + \left( c - \frac{\lambda_p}{2} \right) |V^i_t|^{p+\gamma} + \left( c - \frac{\lambda_p}{2} \right) |\overline{V}^{i,K}_t|^{p+\gamma} \right) |V^i_t - \overline{V}^{i,K}_t|^2 \\
+ C \left( \Lambda_{p+\gamma}(\mu_t^N) + \Lambda_{p+\gamma}(\overline{\mu}^{N,K}_t) \right) d_p^2(t) \\
+ C K^{1-1/v} \left( \Lambda_t(\mu_t^N) + \Lambda_t(\overline{\mu}^{N,K}_t) \right).
$$

Let us now choose $p$. We recall that $c$ does not depend on $p$, and return to the definition

$$
\lambda_p := \int_0^{\pi/2} \left( 1 - \left( 1 + \frac{\cos \theta}{2} \right)^{p/2} \right) \beta(\theta) d\theta.
$$

As $p \to \infty$, the term in parentheses converges up to 1 for any $\theta \neq 0$, and so $\lambda_p$ converges to $\int_0^{\pi/2} \beta(\theta) d\theta = \infty$ by monotone convergence. In particular, there exists some $p_0$, depending only on $G, d$ such that, for all $p > p_0$, $\lambda_p \geq 2c$, and for such $p$, the first line of (104) can be absorbed into the second:

$$
\mathcal{E}_{p,K}(t) \leq C \left( \Lambda_{p+\gamma}(\mu_t^N) + \Lambda_{p+\gamma}(\overline{\mu}^{N,K}_t) \right) d_p^2(t) + C K^{1-1/v} \left( \Lambda_t(\mu_t^N) + \Lambda_t(\overline{\mu}^{N,K}_t) \right)
$$

whence

$$
\overline{d}_p^2(t) \leq \overline{d}_p^2(0) + C \int_0^t \left( \Lambda_{p+\gamma}(\mu_s^N) + \Lambda_{p+\gamma}(\overline{\mu}^{N,K}_s) \right) d_p^2(s) ds \\
+ C K^{1-1/v} \int_0^t \left( \Lambda_t(\mu_s^N) + \Lambda_t(\overline{\mu}^{N,K}_s) \right) ds + \overline{M}_t.
$$
Let us now write \( T := T_b^N \land T_b^{N,K} \) for the stopping times \( T_b^N, T_b^{N,K} \) defined in the statement, and consider the moment prefactor in \((106, 107)\). We recall from Propositions 4, 5 that, almost surely, for all \( t \geq 0 \) and a new choice of \( C / \Lambda_1^p + \gamma (\mu_{t_+}^N) \leq 2 / \Lambda_1^p + \gamma (\tilde{\mu}_{t_+}^{N,K}) \) \((108)\)

and similarly for \( \tilde{\mu}_{t_+}^{N,K} \). The moment factor is therefore at most \( 2b \) for all \( s \leq T \), and so we obtain, for all \( t \geq 0 \),

\[
\int_0^T \left( \Lambda_{p+\gamma}(\mu_s^N) + \Lambda_{p+\gamma}(\tilde{\mu}_s^{N,K}) \right) d\mathbb{P}_s(\cdot) \leq 2b \int_0^T d\mathbb{P}_s(s \land T) ds. \tag{109}
\]

Stopping \((107)\) at \( T \), we therefore obtain, for all \( t \geq 0 \),

\[
\mathbb{E} \left[ \int_0^t \Lambda_{p+\gamma}(\mu_s^N) + \Lambda_{p+\gamma}(\tilde{\mu}_s^{N,K}) \right] ds \leq Cta_2. \tag{111}
\]

We therefore use Grönwall’s Lemma to obtain

\[
\mathbb{E} \left[ \mathbb{E} \left[ \int_0^t \Lambda_{p+\gamma}(\mu_s^N) + \Lambda_{p+\gamma}(\tilde{\mu}_s^{N,K}) \right] ds \right] \leq Cta_2 K^{1-1/\nu}. \tag{112}
\]

Next, we observe that

\[
\mathbb{E} \left[ \int_0^t \Lambda_{p+\gamma}(\mu_s^N) + \Lambda_{p+\gamma}(\tilde{\mu}_s^{N,K}) \right] ds \leq Cta_2 \left( \frac{p+2+\gamma}{p+2+\gamma+\nu} \right) \tag{113}
\]

We now estimate the second term. From the bound \( d_p^2(v, w) \leq c(1 + |v|^{p+2} + |w|^{p+2}) \) we see that

\[
\mathbb{E} \left[ \mathbb{E} \left[ \int_0^t \Lambda_{p+\gamma}(\mu_s^N) + \Lambda_{p+\gamma}(\tilde{\mu}_s^{N,K}) \right] ds \right] \leq C \mathbb{P}(T \leq t)^{1/2} a_3, \tag{115}
\]

thanks to the moment bounds in Lemma 5 and the choice of initial data. Combining with the previous term \((112)\) now proves the first claim.
For the second item, we return to the martingale $\overline{M}_t$ constructed above. From [6, Lemma 8.7], the process

$$L_t = \overline{M}_t^2 - \frac{2}{N^3} \sum_{(i,j)} \int_0^t Q_K (V_i^s, \tilde{V}_i^s, K, V_j^s, \tilde{V}_j^s, K) \, ds$$

(116)
is also a martingale, where the sum now runs over unordered pairs $(i,j)$ of indexes. Thanks to the bound computed in Lemma 6, we find that

$$\mathbb{E} \left[ \overline{M}_{t_{\text{fin}}}^2 \right] \leq \frac{C}{N^3} \mathbb{E} \left[ \sum_{(i,j)} \int_0^{t_{\text{fin}}} (1 + |V_i^s|^q + |V_j^s|^q + |\tilde{V}_i^s|^q + |\tilde{V}_j^s|^q) \, ds \right]$$

$$\leq \frac{C}{N} \int_0^{t_{\text{fin}}} \mathbb{E} (\Lambda_{2/} (\mu_s^N) + \Lambda_{2/} (\tilde{\mu}_s^{N,K})) \, ds.$$ 

(117)

Using the moment propagation estimate in Propositions 4, 5 and Doob’s $L^2$ inequality, we conclude that

$$\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} |\overline{M}_t| \right] \leq \frac{C a_3^{1/2}}{N^{1/2}}. \quad \text{(118)}$$

With this estimate, we return to the argument above. Applying Grönwall to (110), we obtain a pathwise estimate

$$\sup_{t \leq t_{\text{fin}}} \overline{d}_p^2 (t \wedge T) \leq e^{C bt} \left( \overline{d}_p^2 (0) + K^{1-1/y} \int_0^{t_{\text{fin}}} (\Lambda_t (\mu_t^N) + \Lambda_t (\tilde{\mu}_t^{N,K})) \, ds + \sup_{t \leq t_{\text{fin}}} |\overline{M}_t| \right). \quad \text{(119)}$$

Taking expectations, we conclude that

$$\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} \overline{d}_p^2 (t \wedge T) \right] \leq e^{C bt} \left( \overline{d}_p^2 (0) + K^{1-1/y} C t_{\text{fin}} a_2 + \frac{C a_3^{1/2}}{N^{1/2}} \right). \quad \text{(120)}$$

Following the same argument as in (115) we also bound

$$\mathbb{E} \left[ \left( \sup_{t \leq t_{\text{fin}}} \overline{d}_p (t) \right) 1[T \leq t_{\text{fin}}] \right] \leq C \mathbb{P} (T \leq t_{\text{fin}})^{1/2} \mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} (\Lambda_{q} (\mu_t^N) + \Lambda_{q} (\tilde{\mu}_t^{N,K})) \right].$$

[1ex] \leq C \mathbb{P} (T \leq t_{\text{fin}})^{1/2} t_{\text{fin}} a_3. \quad \text{(121)}$$

Combining (120, 121), we obtain the desired result. \qed
6. Proof of Theorem 1

We will now prove Theorem 1, based on the Tanaka coupling presented in Lemma 8 and Corollary 1. As described in Section 1.2 in the introduction, we first take $N \to \infty$ with the cutoff parameters fixed, and then remove the cutoff in a series of lemmata; in order to give an overview of the strategy, we will state all the intermediate steps before turning to the proofs. Our first result transfers the coupling achieved in Corollary 1 to solutions to the cutoff Boltzmann equation, potentially with different cutoff parameters and different initial data.

**Lemma 9.** Let $p > p_0(G, d)$ and $l = p + 2 + \gamma$. Then there exist a constant $C = C(G, p, d)$ such that, whenever $K' \geq K > K_0(G, p, d)$, $a_1, a_2 \geq 1$ and $\mu_0, \nu_0 \in S$ satisfy moment bounds

\[
\Lambda_{p+\gamma}(\mu_0, \nu_0) \leq a_1; \quad \Lambda_l(\mu_0, \nu_0) \leq a_2,
\]

then the solution maps $\phi^K_t$ to the cutoff Boltzmann equation (BE$^K$) satisfy, for all $t \geq 0$,

\[
T_p\left(\phi^K_t(\mu_0), \phi^{K'}_t(\nu_0)\right) \leq e^{C a_1(1+t)} \left( T_p(\mu_0, \nu_0) + a_2 t K^{1-1/\nu}\right).
\]  

(123)

In this way, we take $N \to \infty$, and so the $N$-dependent errors in Lemma 8, which we called type (iii) in the introduction, vanish. As a next step, we show that the solutions $\phi^K_t(\mu_0)$ to the cutoff Boltzmann equations converge, as $K \to \infty$, to a solution of the noncutoff equation (BE).

**Lemma 10.** Let $p, l$ be as above, and let $\mu_0 \in S$ satisfy moment assumptions

\[
\Lambda_{p+\gamma}(\mu_0) \leq a_1, \quad \Lambda_l(\mu_0) \leq a_2
\]

(124)

for some $a_1, a_2 \geq 1$. Then, for some $(\phi_t(\mu))_{t \geq 0} \subset S$ and some $C = C(G, p, d)$,

\[
T_p(\phi^K_t(\mu), \phi_t(\mu)) \leq e^{C a_1(1+t)} t a_2 K^{1-1/\nu}
\]

(125)

for all $K > K_0(G, p, d)$. Moreover, if $\nu_0 \in S$ is another measure with the same moment estimates, we have the continuity

\[
T_p(\phi_t(\mu_0), \phi_t(\nu_0)) \leq e^{C a_1(1+t)} T_p(\mu_0, \nu_0).
\]

(126)

Finally, $(\phi_t(\mu_0) : t \geq 0)$ is a solution to the noncutoff Boltzmann equation (BE), and satisfies the moment estimates in Proposition 5.

The last two statements isolate the errors due to the initial data [type (i) in the terminology of the introduction] and on the cutoff [type (ii)]. Since everything in the type (i) errors depends only on moments of order at most $p + 2$, we can extend the maps $\phi_t$ defined above to all of $S^{p+2+} = \cup_{q>p+2} S^q$, and obtain the claimed continuity estimate in this context.
Lemma 11. Let $p, l$ be as above. The solution maps $\phi_t : S^l \to S$ defined above can be extended to $S^{p+2^+}$, such that, for all $\mu_0 \in S^{p+2^+}$, $(\phi_t(\mu_0) : t \geq 0)$ is a solution to the Boltzmann Equation (BE), and so that (126) holds whenever $\mu_0, \nu_0 \in S^{p+2^+}$ satisfy a moment estimate $\Lambda_{p+\gamma}(\mu_0, \nu_0) \leq a$, for some $a \geq 1$.

To conclude Theorem 1, we must show that the solutions obtained in this way are the unique solutions to (BE) as soon as $\mu_0 \in S^{p+2^+}$. We will use the following auxiliary result, which appears as [17, Corollary 2.3iii]).

Proposition 6. Suppose $\mu_0 \in S$ satisfies, for some $\varepsilon > 0$, $\langle e^{\varepsilon|v|^\gamma}, \mu_0 \rangle < \infty$. Then there exists at most one solution to the Boltzmann Equation (BE) taking values in $S$ and starting at $\mu_0$.

Let us now show how these results imply the claimed result.

Proof of Theorem 1. In light of Lemma 11 above, it remains only to prove that the solutions constructed above are unique. Let us fix $\mu_0 \in S^{p+2^+}$ such that $\Lambda_{p+\gamma}(\mu_0) \leq a$ for some $a \geq 1$. Let $(\mu_t)_{t \geq 0} \subset S$ be any solution to (BE) starting at $\mu_0$; we will now show that $\mu_t = \phi_t(\mu_0)$ for all $t \geq 0$.

Fix $s > 0$, $t \geq 0$. Thanks to the appearance of exponential moments in Proposition 5, there exists $\varepsilon = \varepsilon_s > 0$ such that $\langle e^{\varepsilon|v|^\gamma}, \mu_s \rangle < \infty$, and by Proposition 6, there exists at most one energy-conserving solution starting at $\mu_s$. Since both $(\phi_u(\mu_s))_{u \geq 0}$ and $(\mu_{u+s})_{u \geq 0}$ are such solutions, we conclude that $\phi_t(\mu_s) = \mu_{t+s}$ for all such $t, s$.

Let us now take the limit $s \downarrow 0$. Using the weak formulation of (BE) and the Monge-Kantorovich duality

$$w_1(\mu, \nu) = \sup \left\{ \langle f, \mu - \nu \rangle : \sup_{v \neq w} \frac{|f(v) - f(w)|}{|v - w|} \leq 1 \right\},$$

it is straightforward to see that $w_1(\mu_s, \mu_0) \to 0$, and interpolating between $T_p$ and $w_1$ by (14), it follows that

$$T_p(\mu_s, \mu_0) \leq w_1(\mu_s, \mu_0)^\alpha \Lambda_{p'}(\mu_s, \mu_0)$$

for some $\alpha = \alpha(p') > 0$; the moment factor is bounded uniformly in $s > 0$ by Lemma 5, and so the right-hand side converges to 0 as $s \downarrow 0$. Lemma 11 now shows that, up to a new choice of $C$,

$$T_p(\phi_t(\mu_s), \phi_t(\mu_0)) \leq e^{Ca(1+t)}T_p(\mu_s, \mu_0) \to 0.$$  \hfill (129)

Using the same argument as (128), $T_p(\mu_{t+s}, \mu_t) \to 0$, and we conclude that

$$T_p(\mu_t, \phi_t(\mu_0)) \leq \limsup_{s \downarrow 0} \left[ CT_p(\phi_t(\mu_s), \phi_t(\mu_0)) + CT_p(\mu_{t+s}, \mu_t) \right] = 0$$

and so we have the desired uniqueness. \hfill \Box
6.1. Proof of Lemmata

We will frequently use the following approximation scheme:

**Proposition 7.** [Discrete Approximation Scheme in \( T_p \)] Fix \( p \geq 0 \), and let \( \mu \in S^d \) for \( q > p + 2 \). Then there exists a sequence \( \mu^N \in \mathcal{S}_N \) such that

\[
T_p(\mu^N, \mu) \rightarrow 0; \quad \Lambda_p'(\mu^N) \rightarrow \Lambda_p'(\mu) \quad \text{for all} \quad p' \leq q.
\]

(131)

Thanks to (14), this is equivalent to convergence of the \( q \)th moment and weak convergence, which can easily be obtained by modifying independent random samples; see, for instance, [34, Propositions 9.1 - 9.3].

**Proof of Lemma 9.** Let us consider the case first where the initial data \( \mu_0, \nu_0 \) have a finite \( q \)th moment \( \Lambda_q(\mu_0, \nu_0) \leq a_3 \) for some \( a_3 \geq 1 \). Applying Proposition 7, take \( N \)-particle empirical measures \( \mu^N_0 \in \mathcal{S}_N \) such that \( T_p(\mu^N_0, \mu_0) \rightarrow 0 \) and such that the \( l \)th, \( q \)th moments converge: \( \Lambda_l(\mu^N_0) \rightarrow \Lambda_l(\mu_0) \), \( \Lambda_q(\mu^N_0) \rightarrow \Lambda_q(\mu_0) \); construct \( \nu^N_0 \) similarly for \( \nu_0 \). Using the relaxed triangle inequality (12), it follows that, for some \( C = C(p) \),

\[
\lim_{N \rightarrow \infty} \sup_{p \geq 0} T_p(\mu^N_0, \nu^N_0) \leq C \cdot T_p(\mu_0, \nu_0).
\]

(132)

Let us now take \( \mu^N_t, \nu^N_t \) to be cutoff Kac processes given by Corollary 1ii), with cutoff parameters \( K, K' \), respectively, and started at these initial data; fix \( t \geq 0 \), and consider

\[
T_p(\phi^K_t(\mu_0), \phi^{K'}_t(\nu_0)) \leq C \mathbb{E} \left[ T_p(\phi^K_t(\mu_0), \mu^N_t, \nu^N_t) + T_p(\nu^N_t, \phi^{K'}_t(\nu_0)) \right].
\]

(133)

The construction via Corollary 1ii) bounds the middle term, giving

\[
T_p(\phi^K_t(\mu_0), \phi^{K'}_t(\nu_0)) \leq C \mathbb{E} \left[ T_p(\phi^K_t(\mu_0), \mu^N_t, \nu^N_t) + T_p(\nu^N_t, \phi^{K'}_t(\nu_0)) \right] + e^{C \alpha_1(1+t)} \left( T_p(\mu^N_0, \nu^N_0) + a_2 K^{1-1/v} \right) + a_3^2 C N^{-1/2}.
\]

(134)

Let us now take \( N \rightarrow \infty \). Thanks to Lemma 2, both terms on the first line converge to 0, as does the term on the third line. Using (132), we conclude that

\[
T_p(\phi^K_t(\mu_0), \phi^{K'}_t(\nu_0)) \leq e^{C \alpha_1(1+t)} \left( T_p(\mu^N_0, \nu^N_0) + a_2 K^{1-1/v} \right)
\]

(135)

as claimed.

Let us now show how this extends to initial data \( \mu_0, \nu_0 \) with only \( l = p + \gamma + 2 \) moments as in the statement. In this case, we use Proposition 7 again, with \( l \) in place of \( q \), to construct \( \mu^N_0 \in \mathcal{S}_N \) such that

\[
T_p(\mu^N_0, \mu_0) \rightarrow 0; \quad \Lambda_l(\mu^N_0) \rightarrow \Lambda_l(\mu_0),
\]

(136)
and similarly \( v_0^N \) for \( v_0 \). Since \( \mu_0^N, v_0^N \) are compactly supported, the previous estimate applies so that

\[
T_p \left( \phi^K_0(\mu^0), \phi^K(v_0^N) \right) \leq e^{C_1(1+t)} \left( T_p \left( \mu^0, v_0^N \right) + a_2 K^{1-1/v} \right).
\]  

(137)

Using Lemma 3,

\[
T_p \left( \phi^K(\mu^0), \phi^K(\mu_0) \right) \rightarrow 0; \quad T_p \left( \phi^K(v_0^N), \phi^K(v_0) \right) \rightarrow 0.
\]  

(138)

The same argument as above therefore allows us to take \( N \rightarrow \infty \) in (137), noting that no moments higher than \( l \)th appear, to conclude that

\[
T_p \left( \phi^K(\mu_0), \phi^K(v_0) \right) \leq C e^{C_1(1+t)} \left( T_p \left( \mu_0, v_0 \right) + a_2 K^{1-1/v} \right).
\]  

(139)

\[\square\]

Proof of Lemma 10. Let us consider the space \( S_{\leq 1} \) of probability measures \( \mu \) on \( \mathbb{R}^d \) satisfying \( \langle |v|^2, \mu \rangle \leq 1 \), which we equip with the Wasserstein_1-distance \( w_1 \); it is straightforward to see that this space is complete. We now define

\[
C := C([0, \infty), (S_{\leq 1}, w_1))
\]  

(140)

equipped with a metric inducing uniform convergence on compact time intervals; since \( (S_{\leq 1}, w_1) \) is complete, so is \( C \). It is straightforward to see that \( w_1 \leq T_p \), and so the previous observation shows that \( (\phi^K(\mu_0), t \geq 0)_{K \geq 1} \) are Cauchy in \( C \) for any fixed \( \mu_0 \in S' \), and hence converge to some \( (\phi_t(\mu_0), t \geq 0) \in C \).

Next, let us show that \( \phi^K_t(\mu_0) \rightarrow \phi_t(\mu_0) \) in \( T_p \). For \( t = 0 \), \( \phi^K_0(\mu_0) = \mu_0 \), and so there is nothing to prove. If \( t > 0 \) then, by point iii) of Proposition 4, there exists \( \lambda_{p+3} = \lambda_{p+3}(t) < \infty \) such that, for all \( K \geq 1 \),

\[
\Lambda_{p+3} \left( \phi^K_t(\mu_0) \right) \leq \lambda_{p+3}(t).
\]  

(141)

Using the boundedness of these \( p+3 > 2 \) moments, it follows that \( 0 = \langle v, \phi^K_t(\mu_0) \rangle \rightarrow \langle v, \phi_t(\mu_0) \rangle, 1 = \langle |v|^2, \phi^K_t(\mu_0) \rangle \rightarrow \langle |v|^2, \phi_t(\mu_0) \rangle \), and so the limit \( \phi_t(\mu_0) \) lies in \( S \). By lower semicontinuity of the moments \( \mu \mapsto \Lambda_{p+3}(\mu) \) in \( w_1 \), the same is true for the limit \( \phi_t(\mu_0) \), and interpolating between \( w_1 \) and \( T_p \) by (14) gives

\[
T_p \left( \phi^K_t(\mu_0), \phi_t(\mu_0) \right) \leq \Lambda_{p+3} \left( \phi^K_t(\mu_0), \phi_t(\mu_0) \right) w_1 \left( \phi^K_t(\mu_0), \phi_t(\mu_0) \right)^\alpha
\]  

(142)

for some \( \alpha > 0 \). By construction, the second term on the right-hand side converges to 0, and the first term is bounded, so the left-hand side converges to 0 as desired.

We now conclude the bound (125): if \( K > K_0(G, p, d) \), then for all \( K' \geq K \),

\[
T_p \left( \phi^K_t(\mu_0), \phi_t(\mu_0) \right) \leq C \left( T_p \left( \phi^K_t(\mu_0), \phi^K_t(\mu_0) \right) + T_p \left( \phi^K_t(\mu_0), \phi_t(\mu_0) \right) \right)
\]

\[
\leq C \left( e^{C_1(1+t)} a_2 K^{1-1/v} + T_p \left( \phi^K_t(\mu_0), \phi_t(\mu_0) \right) \right).
\]  

(143)
Taking $K' \to \infty$, the second term on the final line converges to 0, and the desired bound follows, absorbing the prefactor $C = C(p)$ into the exponent. The bound (126) is similar: if $\mu_0, v_0$ in $S^l$ satisfy
\[ \Lambda_{p+\gamma}(\mu_0, v_0) \leq a_1; \quad \Lambda_{l}(\mu_0, v_0) \leq a_2 \] (144)
then we bound, for any $K$,
\[
T_p(\phi_t(\mu_0), \phi_t(v_0)) \leq C(T_p\left(\phi^K_t(\mu_0), \phi_t(\mu_0)\right) + T_p\left(\phi^K_t(v_0), \phi_t(v_0)\right) \quad (145)
\]
where, in the second line, we have used Lemma 9 to compare $\phi^K_t(\mu_0), \phi^K_t(v_0)$ and used the previous part to estimate the other two terms. Taking $K \to \infty$, we conclude the desired bound, again up to a new choice of $C$.

It remains to show that $\phi_t(\mu_0), t \geq 0$ solves the full, noncutoff Boltzmann equation (BE). We begin with an analysis of the the Boltzmann collision operator, borrowing from [30]. Let us define, for bounded, Lipschitz $f : \mathbb{R}^d \to \mathbb{R}$,
\[
(\Delta_B f)(v, v_*) := \int_{S^{d-1}} (f(v') + f(v_*') - f(v) - f(v_*))B(v - v_*, \sigma) (146)
\]
and observe that
\[
\langle f, Q(\mu) \rangle = \int_\mathbb{R}^d \times \mathbb{R}^d (\Delta_B f)(v, v_*) \mu(\text{d}v) \mu(\text{d}v_*) \quad (148)
\]
and similarly for $B_K$. It is straightforward to see that each $\Delta_{B_K} f$ is continuous on $\mathbb{R}^d \times \mathbb{R}^d$, and the straightforward estimate $|v' - v| \leq |v - v_*| \sin \theta$ implies the growth bound
\[
|(\Delta_B f)(v, v_*)| \leq C(f)|v - v_*| \int_{S^{d-1}} \sin \theta B(v - v_*, \sigma) d\sigma \leq C(f)|v - v_*|^{1+\gamma} \quad (149)
\]
for some constant $C = C(f)$, depending only on the Lipschitz constant of $f$, and similarly for $B_K$. The same argument also shows that
\[
|(\Delta_B f)(v, v_*) - (\Delta_{B_K} f)(v, v_*)| \leq C(f)\varepsilon_K |v - v_*|^{1+\gamma}; \quad (150)
\]
\[
\varepsilon_K = \int_{S^{d-1}} (\sin \theta) 1(\theta \leq \theta_0(K)) B(u, \text{d}\sigma) \to 0 \quad (151)
\]
so that $\Delta_{B_K} f \to \Delta f$, uniformly on compact subsets of $\mathbb{R}^d \times \mathbb{R}^d$; it therefore follows that $\Delta_B f$ is continuous.
Equipped with this dual formulation, let us fix $t \geq 0$ and a bounded, Lipschitz $f$. Writing $\mu^K_t := \phi^K_t(\mu_0), \mu_t := \phi_t(\mu_0)$, we claim that
\[
(f, Q(K(\mu^K_t))) \to (f, Q(\mu_t)) \tag{152}
\]
for any $t \geq 0$. For all $R \geq 0$, let $\psi_R : \mathbb{R}^d \times \mathbb{R}^d \to [0, 1]$ be a smooth, compactly supported cutoff function, such that $\psi_R(v, v_*) = 1$ on the ball $|v|^2 + |v_*|^2 \leq R$. We estimate, uniformly in $K$,
\[
\int_{\mathbb{R}^d \times \mathbb{R}^d} |(\Delta_B f)(v, v_*)| \mu^K_t(dv) \mu^K_t(dv_*) \\
\leq C(f) \int_{\mathbb{R}^d \times \mathbb{R}^d} (1 + |v|^2 + |v_*|^2) 1(|v|^2 + |v_*|^2 \geq R) \mu^K_t(dv) \mu^K_t(dv_*) \tag{153}
\]
where, in the final line, we used the moment hypothesis on $\mu_0$, with $l > p + 2$, and the moment propagation result in Proposition 4; the same argument holds for the limit $\mu_0$ implies that, for all compactly supported, continuous $g$ and, combining with (157), we see that
\[
\psi \text{ supported cutoff function, such that } \\
\int_{\mathbb{R}^d \times \mathbb{R}^d} g(v, v_*) \left( \mu^K_t(dv) \mu^K_t(dv_*) - \mu_t(dv) \mu_t(\mu_0)(dv_*) \right) \to 0, \tag{154}
\]
and, in particular, this holds with $g = (\Delta_B f)(v, v_*)\psi_R$. We now write
\[
\left| \left( f, Q(\mu_t) - Q(\mu^K_t) \right) \right| \\
\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |(\Delta_B f)| (1 - \psi_R)(v, v_*) (\mu^K_t(dv) \mu^K_t(dv_*) + \mu_t(dv) \mu_t(dv_*)) + \int_{\mathbb{R}^d \times \mathbb{R}^d} (\Delta_B f) \psi_R(v, v_*) (\mu^K_t(dv) \mu^K_t(dv_*) - \mu_t(dv) \mu_t(dv_*)) \tag{155}
\]
The second term converges to 0 by (154), so using (153) twice on the first term,
\[
\limsup_{K \to \infty} \left| \left( f, Q(\mu_t) - Q(\mu^K_t) \right) \right| \leq C R^{-p} a_2 \Lambda_l(\mu_0) \tag{156}
\]
and, since $R$ was arbitrary, we have shown that
\[
\left( f, Q(\mu^K_t) \right) \to (f, Q(\mu_t)) . \tag{157}
\]
Finally, integrating (150), we find
\[
\left| \left( f, Q(\mu_t^K) - Q_K(\mu^K_t) \right) \right| \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |(\Delta_B f) - (\Delta_{B_K} f)|(v, v_*) \mu^K_t(dv) \mu^K_t(dv_*) \\
\leq C(f) \varepsilon_K \int_{\mathbb{R}^d \times \mathbb{R}^d} (1 + |v|^2 + |v_*|^2) \mu^K_t(dv) \mu^K_t(dv_*) \tag{158}
\]
and, combining with (157), we see that $(f, Q(\mu^K_t)) \to (f, Q(\mu_t))$ as claimed.
We now conclude. For any $t \geq 0$ and any bounded, Lipschitz $f$, we have
\begin{equation}
\langle f, \mu^K_t \rangle = \langle f, \mu_0 \rangle + \int_0^t \langle f, Q_K(\mu^K_s) \rangle ds.
\end{equation}

The integrand $\langle f, Q_K(\mu^K_s) \rangle$ is bounded, uniformly in $s \leq t$ and $K \geq 1$, and converges to $\langle f, Q(\mu_s) \rangle$ for all $s$, while the left-hand side converges to $\langle f, \mu_t \rangle$. We therefore take the limit $K \to \infty$ to conclude that, for all bounded, Lipschitz $f$ and all $t \geq 0$
\begin{equation}
\langle f, \mu_t \rangle = \langle f, \mu_0 \rangle + \int_0^t \langle f, Q(\mu_s) \rangle ds
\end{equation}
as desired. \hfill \Box

Finally, Lemma 11 follows much the same pattern as above.

**Proof of Lemma 11.** To extend the maps $\phi_t$, fix $\mu_0 \in S^{p+2+}$. Using Proposition 7 again, let $\mu_0^N \in S_N$ be a sequence of discrete measures such that $T_p(\mu_0^N, \mu_0) \to 0$ and $\Lambda_{p+\gamma}(\mu_0^N) \to \Lambda_{p+\gamma}(\mu_0)$; in particular, $\Lambda_{p+\gamma}(\mu_0^N) \leq 2a$ for all $N$ large enough. The bound (126) obtained in the previous lemma applies to show that, for all such $N$ and all $t \geq 0$,
\begin{equation}
w_1(\phi_t(\mu_0^N), \phi_t(\mu_0^{N'})) \leq T_p \left( \phi_t(\mu_0^N), \phi_t(\mu_0^{N'}) \right) \leq e^{2Ca(1+t)} T_p(\mu_0^N, \mu_0^{N'}).\end{equation}

The right-hand side converges to 0 as $N, N' \to \infty$, which implies that $\phi_t(\mu_0^N)$ converges, uniformly in $w_1$ on compact time intervals, to some limit. If we now define $\phi_t(\mu_0)$ to be this limit, a similar calculation shows that the resulting $\phi_t(\mu_0)$ is independent of the choice of limiting sequence, and the same argument as in Lemma 10 above shows that $(\phi_t(\mu_0), t \geq 0) \subset S$ is again a solution to the noncutoff Boltzmann equation (BE). Finally, if $\mu_0, v_0$ are two such measures, one applies (126) to approximating sequences $\mu_0^N, v_0^N$ and passes to the limit $N \to \infty$ to obtain the same result for $\mu_0, v_0$, again up to a new constant $C$ in the exponent.

### 7. Proof of Theorem 2

We now prove the Theorem 2 concerning the convergence of the full, non-cutoff Kac process to the solution to the Boltzmann equation in the many-particle limit $N \to \infty$.

**Proof of Theorem 2.** The uniqueness in law follows from Propositions 1, 3, which are discussed in Appendix B.

For the convergence estimate, let $\mu^N_t, t \geq 0$ be any unlabelled Kac process, and consider the case $\mu_0 = \mu_0^N$. Fix $t_{\text{fin}}$ and $K$ to be chosen later; for this $K$, let $\tilde{\mu}_t^N, \tilde{\mu}_t^{N,K}$ be the coupling of noncutoff and cutoff Kac processes, both starting at $\mu^N_0$, given by Corollary 1. By uniqueness in law, it is sufficient to prove the estimate
with $\tilde{\mu}_i^N$ in place of $\mu_i^N$. For some constants $C = C(p, q), \alpha = \alpha(p, q)$, we have the following estimates. By Corollary 1,

$$
E \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \tilde{\mu}_i^N, \tilde{\mu}_i^N, K \right) \right] \leq e^{Ca(1 + t_{\text{fin}})} (K^{1-1/\nu} + N^{-1/2}); \tag{162}
$$

by Lemma 2

$$
E \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \tilde{\mu}_i^N, K, \phi_t^K (\mu_0^N) \right) \right] \leq \exp (CaK (1 + t_{\text{fin}})) N^{-\alpha}; \tag{163}
$$

and by Lemma 10,

$$
\sup_{t \leq t_{\text{fin}}} T_p \left( \phi_t^K (\mu_0^N), \phi_t (\mu_0^N) \right) \leq e^{Ca(1 + t_{\text{fin}})} K^{-1/\nu}. \tag{164}
$$

Combining, and keeping the worst terms, we have the estimate

$$
E \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \tilde{\mu}_i^N, \phi_t (\mu_0^N) \right) \right] \leq e^{Ca(1 + t_{\text{fin}})} K^{1-1/\nu} + e^{CaK(1 + t_{\text{fin}})} N^{-\alpha}. \tag{165}
$$

We now choose

$$
K = \max \left( 1, \frac{1}{2Ca(1 + t_{\text{fin}})} \log (N^\alpha) \right) \tag{166}
$$

to conclude that

$$
E \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \tilde{\mu}_i^N, \phi_t (\mu_0^N) \right) \right] \leq e^{Ca(1 + t_{\text{fin}})} (\log N)^{1-1/\nu}. \tag{167}
$$

Finally, by Theorem 1, we have

$$
E \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \phi_t (\mu_0^N), \phi_t (\mu_0) \right) \right] \leq e^{Ca(1 + t_{\text{fin}})} E \left[ T_p (\mu_0^N, \mu_0) \right], \tag{168}
$$

and combining gives the claimed bound. \hfill \Box

8. Main Calculations on the Tanaka Coupling

8.1. Some estimates for $G$

In preparation for the proofs of Lemma 5, we will first record some basic estimates concerning the regularity and integrability of $G$.  

Lemma 12. (i) Let $G$ be as above. Then, for some constants $0 < c_1 \leq c_2 < \infty$, we have

$$c_1 (1 + z)^{-1/\nu} \leq G(z) \leq c_2 (1 + z)^{-1/\nu}. \quad (169)$$

Moreover, $G$ is continuously differentiable, and $c_1, c_2$ above can be chosen such that

$$c_1 (1 + z)^{-1-1/\nu} \leq G'(z) \leq c_2 (1 + z)^{-1-1/\nu}. \quad (170)$$

(ii) We have

$$\int_0^\infty \left| \frac{d}{dz} (z(1 - \cos G(z))) \right| \, dz < \infty. \quad (171)$$

(iii) There exists a constant $c < \infty$ such that, for all $x, y > 0$,

$$\int_0^\infty \left( G \left( \frac{z}{x} \right) - G \left( \frac{z}{y} \right) \right)^2 \, dz \leq c \frac{|x - y|^2}{x + y}. \quad (172)$$

Proof. (i) For the first claim, we use the definition of $H$ and (3) to see that, for some constants $c_1, c_2 \in (0, \infty)$ and all $\theta \in (0, \pi/2)$,

$$c \int_\theta^{\pi/2} x^{-1-\nu} \, dx \leq H(\theta) \leq C \int_\theta^{\pi/2} x^{-1-\nu} \, dx \quad (173)$$

so that

$$\frac{c_1}{\nu} \left( \theta^{-\nu} - \left( \frac{\pi}{2} \right)^{-\nu} \right) \leq H(\theta) \leq \frac{c_2}{\nu} \left( \theta^{-\nu} - \left( \frac{\pi}{2} \right)^{-\nu} \right). \quad (174)$$

The first claim now follows, potentially for a new choice of $c_1, c_2$. The differentiability is an immediate consequence of the inverse function theorem. Indeed, we have

$$G'(z) = \frac{1}{H'(G(z))} = -\frac{1}{b(\cos G(z))} \quad (175)$$

and so the second claim follows from the first, using (3).

(ii) We expand the derivative

$$\frac{d}{dz} (z(1 - \cos G(z))) = (1 - \cos G(z)) + z (\sin G(z)) \, G'(z). \quad (176)$$

Therefore, we have

$$\left| \frac{d}{dz} (z(1 - \cos G(z))) \right| \leq G(z)^2 + zG(z)|G'(z)|. \quad (177)$$

Using the bounds from the previous part, it follows that the right-hand side is bounded by $c_2 (1 + z)^{-2/\nu}$ for some $c_2 < \infty$, which is integrable because $\nu \in (0, 1)$. 
The following is a slight variant of [16, Lemma 1.1], and is included here for completeness. Recalling that $G$ is decreasing, and integrating the bound on $G'$ found in part i), we see that, for all $0 \leq z \leq w$ and some $c < \infty$, we have

$$0 \leq G(z) - G(w) \leq c \left( (1+z)^{-1/v} - (1+w)^{-1/v} \right).$$

We also recall that, for all $a > b > 0$, we have

$$a^{1/v} - b^{1/v} \leq c \frac{a - b}{a^{1-1/v} + b^{1-1/v}}.$$  \hspace{1cm} (179)

For any $z > 0, 0 < y < x$, we apply this bound with $a = (1+z/x)^{-1}, b = (1+z/y)^{-1}$ to obtain

$$0 \leq G \left( \frac{z}{x} \right) - G \left( \frac{z}{y} \right) \leq c \left( (1+z/x)^{-1/v} - (1+z/y)^{-1/v} \right) \leq c \left| \frac{x}{x+z} - \frac{y}{y+z} \right| \left( 1 + \frac{z}{x} \right)^{1-1/v} \leq c |x - y|(x + z)^{-1/v} x^{-1+1/v}. \hspace{1cm} (180)$$

We square and integrate over $z$, to obtain for all $x > y > 0$,

$$\int_0^\infty \left( G \left( \frac{z}{x} \right) - G \left( \frac{z}{y} \right) \right)^2 \, dz \leq c |x - y|^2 x^{1-2/v} x^{-2+2/v} = c \frac{|x - y|^2}{x + y}. \hspace{1cm} (181)$$

This concludes the proof of both claimed bounds in the case $x > y > 0$; for $y > x$, we reverse the roles of $x \leftrightarrow y$. \hspace{1cm} \square

8.2. Proof of Lemma 5

We now turn to the proof of Lemma 5, which was deferred earlier. In order to avoid unnecessarily unwieldy expressions, we introduce some notation. We define $x = |v - v_*|, \tilde{x} = |\nu - \tilde{\nu}|$, and write $L$ for the cutoff $L = K_{\tilde{\nu}} \nu$. We will also write $R$ for $R(v - v_*, \nu - \tilde{\nu}_*)$, and suppress the dependence of $a, \tilde{a}_K, \mathcal{E}_K$ on $v, \nu, v_*, \tilde{\nu}_*$. Throughout, $c$ will denote a constant which is allowed to depend only on $G$, $d$, and $C$ will denote a constant which is also allowed to depend on $p$; both are understood to vary from line to line as necessary.

In order to make the calculations slightly more palatable, we will break up $\mathcal{E}_{p, K}$ as follows. We define

$$\mathcal{E}^1_{p, K} = \int_0^\infty \, dz \int_{\mathbb{S}^{d-2}} \, d\varphi \left( |v'|^p |v' - \tilde{\nu}'_K|^2 - |v|^p |v - \tilde{\nu}|^2 \right); \hspace{1cm} (182)$$

$$\mathcal{E}^2_{p, K}(v, \nu, v_*, \tilde{\nu}_*) = \int_0^\infty \, dz \int_{\mathbb{S}^{d-2}} \, d\varphi \left( |\nu_K' |p |v' - \tilde{\nu}'_K|^2 - |\nu|^p |v - \tilde{\nu}|^2 \right); \hspace{1cm} (183)$$

$$\mathcal{E}^3_{p, K}(v, \nu, v_*, \tilde{\nu}_*) = \int_0^\infty \, dz \int_{\mathbb{S}^{d-2}} \, d\varphi \left( |v' - \tilde{\nu}'_K|^2 - |v - \tilde{\nu}|^2 \right). \hspace{1cm} (184)$$
In this way, using the definition of \(d_p\), it follows that \(\mathcal{E}_{p,K} = \mathcal{E}^1_{p,K} + \mathcal{E}^2_{p,K} + \mathcal{E}^3_{p,K}\). It therefore suffices to prove the following estimates.

**Lemma 13.** For some constants \(K_0 = K_0(p), c = c(G, d)\) and \(C = C(G, d, p)\), and \(q = p + 2 + \gamma\), whenever \(K \geq K_0(p)\), we have

\[
\mathcal{E}^1_{p,K}(v, \tilde{v}, v_*, \tilde{v}_*) \leq \left( c + \left( c - \frac{\lambda_p}{2} \right) |v|^{p+\gamma} + c |\tilde{v}|^{p+\gamma} \right) |v - \tilde{v}|^2 \\
+ (c|v_*|^{p+\gamma} + c|\tilde{v}_*|^{p+\gamma}) |v_* - \tilde{v}_*|^2 \\
+ C \left( (|v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma})(1 + |v|^{p} + |\tilde{v}|^{p}) |v - \tilde{v}|^2 \\
+ C \left( (|v|^{p+\gamma} + |\tilde{v}|^{p+\gamma})(1 + |v_*|^{p} + |\tilde{v}_*|^{p}) |v_* - \tilde{v}_*|^2 \\
+ cK^{1-1/\nu}(1 + |v|^l + |v_*|^l + |\tilde{v}|^l + |\tilde{v}_*|^l) \right) \right)
\quad (185)
\]

and

\[
\mathcal{E}^3_{p,K}(v, \tilde{v}, v_*, \tilde{v}_*) \leq c(1 + |v|^{\gamma} + |\tilde{v}|^{\gamma} + |v_*|^{\gamma} + |\tilde{v}_*|^{\gamma})(|v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2) \\
+ cK^{1-1/\nu}(1 + |v|^l + |\tilde{v}|^l + |v_*|^l + |\tilde{v}_*|^l). \quad (187)
\]

**Proof of Lemmata 5, 13.** Let us begin from the bound, for any \(p \geq 4\), obtained at (36),

\[
|v + a|^p \leq |v|^p \left( \frac{1 + \cos G(z/x^{\gamma})}{2} \right)^{p/2} + |v_*|^p \left( \frac{\sin G(z/x^{\gamma})}{2} \right)^{p/2} \\
+ C \left( |v|^{p-2}|v_*|^2 + |v|^2|v_*|^{p-2} \right) \sin G(z/x^{\gamma}) \\
=: f_p(|v|, |v_*|, z, x).
\]

Let us also write

\[
f_p^*(|v|, |v_*|) = |v|^p + C(|v|^2|v_*|^{p-2} + |v|^p|v_*|^{p-2}) + |v_*|^p \quad (189)
\]

which is an upper bound for \(f_p\), uniformly in \(z, x\). We therefore find

\[
\mathcal{E}^1_{p,K} \leq \int_0^\infty dz \int_{S^{d-2}} d\varphi \left( f_p(|v|, |v_*|, z, x)|v - \tilde{v} + a - \tilde{a}|^2 - |v|^p|v - \tilde{v}|^2 \right).
\quad (190)
\]

Let us also introduce

\[
\tilde{a} = a(\tilde{v}, \tilde{v}_*, z, R(v - v_*, \tilde{v} - \tilde{v}_*)\varphi)
\quad (191)
\]
so that \( \tilde{a} = a_1(z \leq L) \). We can therefore replace \( \tilde{a} \) by \( a \), introducing a further error:

\[
\mathcal{E}_{p,K}^1 \leq \int_0^\infty dz \int_{S^{d-2}} d\varphi \ (f_p(|v|, |v_s|, z, x)|v - \tilde{v} + a - \tilde{a}|^2 - |v|^p|v - \tilde{v}|^2)
+ \int_L^\infty dz \int_{S^{d-2}} d\varphi \ f_p(|v|, |v_s|, z, x) \left(|v - \tilde{v} + a|^2 - |v - \tilde{v} + a - \tilde{a}|^2\right). \tag{192}
\]

Finally, we expand the squared norm \(|v - \tilde{v} + a - \tilde{a}|^2\) in the first line to obtain the decomposition

\[
\mathcal{E}_{p,K}^1 \leq H_1 + H_2 + H_3 + H_4, \tag{193}
\]

where we define

\[
H_1 := \int_0^\infty dz \int_{S^{d-2}} d\varphi \ (f_p(|v|, |v_s|, z, x) - |v|^p)|v - \tilde{v}|^2; \tag{194}
\]

\[
H_2 := 2 \int_0^\infty dz \int_{S^{d-2}} d\varphi \ f_p(|v|, |v_s|, z, x)(v - \tilde{v}) \cdot (a - \tilde{a}); \tag{195}
\]

\[
H_3 := \int_0^\infty dz \int_{S^{d-2}} d\varphi \ f_p^*(|v|, |v_s|)|a - \tilde{a}|^2; \tag{196}
\]

\[
H_4 := \int_L^\infty dz \int_{S^{d-2}} d\varphi \ f_p(|v|, |v_s|, z, x) \left(|v + a - \tilde{v}|^2 - |v + a - \tilde{v} - \tilde{a}|^2\right). \tag{197}
\]

We will now analyse this bound for \( \mathcal{E}_{p,K}^1 \) in detail, and an equivalent analysis of \( \mathcal{E}_{p,K}^2, \mathcal{E}_{p,K}^3 \) will be discussed at the end of the proof. Let us now deal with these terms one by one.

1. Analysis of \( H_1 \). Recalling the construction of \( G \), the moment integral in \( H_1 \) can be reparametrised in terms of \( \theta \):

\[
\int_0^\infty (f_p(|v|, |v_s|, z, x) - |v|^p)dz
= -|v - v_s|^\gamma |v|^p \int_0^{\pi/2} \beta(\theta) \left(1 - \left(\frac{1 + \cos \theta}{2}\right)^{p/2}\right) d\theta
+ |v - v_s|^\gamma (|v_s|^p + C(|v|^2|v_s|^{p-2} + |v|^{p-2}|v_s|^2)) \int_0^{\pi/2} \beta(\theta) \sin(\theta) d\theta
\leq -\lambda_p |v - v_s|^\gamma |v|^p + C|v - v_s|^\gamma (|v_s|^p + |v|^2|v_s|^{p-2} + |v|^{p-2}|v_s|^2).
\]

We now use the bound \(|v|^\gamma - |v_s|^\gamma \leq |v - v_s|^\gamma \leq |v|^\gamma + |v_s|^\gamma\) and the Peter-Paul inequality to obtain

\[
\int_0^\infty (f_p(|v|, |v_s|, z, x) - |v|^p)dz \leq -\frac{\lambda_p}{2} |v|^{p+\gamma} + C|v_s|^{p+\gamma} \tag{199}
\]

and so

\[
H_1 \leq -\frac{\lambda_p}{2} |v|^{p+\gamma} |v - \tilde{v}|^2 + C|v_s|^{p+\gamma} |v - \tilde{v}|^2. \tag{200}
\]
2. Analysis of $H_2$. We first observe that

$$
\int_{S^{d-2}} d\varphi \ (a - \bar{a}) = -\frac{1}{2} (1 - \cos(G(x/y'))) (v - v_\star) + \frac{1}{2} (1 - \cos(G(z/\bar{x}'))) (\bar{v} - \bar{v}_\star).
$$

(201)

It therefore follows that

$$
H_2 = (v - \bar{v}) \cdot \{ \Phi(\bar{x}, |v|, |v_\star|, x)(\bar{v} - \bar{v}_\star) - \Phi(x, |v|, |v_\star|, x)(v - v_\star) \}
$$

(202)

where we define, for any $y, u, v, w > 0$,

$$
\Phi(y, u, v, w) = \int_0^\infty dz \ f_p(u, v, z, w)(1 - \cos(G(z/y'))) \ dz
:= \Psi(y', u, v, w).
$$

(203)

We now rewrite $\Psi$ as

$$
\Psi(y, u, v, w) = y \int_0^\infty dz \ f_p(u, v, zy, w)(1 - \cos(G(z))) \ dz.
$$

(204)

Written in this form, it is clear that $\Psi$ is differentiable, and we calculate

$$
\frac{\partial}{\partial y} \Psi(y, u, v, w) = \int_0^\infty f_p(u, v, zy, w)(1 - \cos(G(z))) \ dz + y \int_0^\infty (1 - \cos(G(z))) \left( \frac{z}{y} \right) \left( \frac{\partial}{\partial z} f_p(u, v, zy, w) \right) \ dz
$$

(205)

$$
= \int_0^\infty f_p(u, v, zx, w)(1 - \cos(G(z))) \ dz - \int_0^\infty \left( \frac{d}{dz} \left( z(1 - \cos(G(z)) \right) \right) f_p(u, v, zx, w) \ dz
$$

where the final line follows by an integration by parts. From the calculations in Lemma 12, we therefore conclude that

$$
\left| \frac{\partial}{\partial y} \Psi(y, u, v, w) \right| \leq c f_p^*(u, v).
$$

(206)

Now, using the bound $|x'^y - y'^y| \leq 2|x - y|/(x^{1 - \gamma} + y^{1 - \gamma})$, we obtain

$$
|\Phi(x, |v|, |v_\star|, x) - \Phi(\bar{x}, |v|, |v_\star|, x)| \leq \frac{c |x - \bar{x}|}{x^{1 - \gamma} + \bar{x}^{1 - \gamma}} f_p^* (|v|, |v_\star|)
$$

(207)

and, for all $y > 0$,

$$
|\Phi(x, |v|, |v_\star|, y)| \leq cy'^y f_p^* (|v|, |v_\star|).
$$

(208)
We therefore obtain the bound
\[ |H_2| \leq |v - \tilde{v}| \left\{ |v - v_* - \tilde{v} + \tilde{v}_*| |\Phi(x, |v|, |v_*|, x) + \Phi(\tilde{x}, |v|, |v_*|, x)| + (|v - v_*| + |\tilde{v} - \tilde{v}_*|) |\Phi(x, |v|, |v_*|, x) - \Phi(\tilde{x}, |v|, |v_*|, x)| \right\} \tag{209} \]
\[ \leq c \left( |v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2 \right) (|v|^2 + |u_*|^2 + |\tilde{v}|^2 + |\tilde{v}_*|^2) \]

3. Analysis of $H_3$. We now turn to the term $H_3$, and begin by noting that
\[ a \cdot \tilde{a} = \frac{1}{4} \left( 1 - \cos G(z/x^\gamma) \right) \left( 1 - \cos G(z/\tilde{x}^\gamma) \right) (v - v_*) \cdot (\tilde{v} - \tilde{v}_*) \]
\[ - \frac{1}{4} \left( 1 - \cos G(z/x^\gamma) \right) \sin G(z/\tilde{x}^\gamma)(v - v_*) \cdot (\tilde{v} - \tilde{v}_*) \]
\[ - \frac{1}{4} \left( 1 - \cos G(z/\tilde{x}^\gamma) \right) \sin G(z/x^\gamma) \Gamma(v - v_*, \varphi) \cdot (\tilde{v} - \tilde{v}_*) \]
\[ + \frac{1}{4} \sin G(z/x^\gamma) \sin G(z/\tilde{x}^\gamma) \Gamma^2(v - v_*, \varphi) \cdot (\tilde{v} - \tilde{v}_*, \varphi) \tag{210} \]

We now integrate over $\varphi \in S^{d-2}$. Since $\int_{S^{d-2}} \Gamma(u, \varphi) d\varphi = 0$ and $R$ preserves the uniform measure $d\varphi$, the middle two lines integrate to 0. We also recall, from the construction of $R = R(v - v_*, \tilde{v} - \tilde{v}_*)$ in Lemma 4, that $\Gamma(v - v_*, \varphi) \cdot (\tilde{v} - \tilde{v}_*, \varphi) \geq (v - v_*, \tilde{v} - \tilde{v}_*)$, and so integrating (210) gives
\[ \int_{S^{d-2}} a \cdot \tilde{a} \ d\varphi \leq \frac{1}{4} \left[ (1 - \cos G(z/x^\gamma)) \left( 1 - \cos G(z/\tilde{x}^\gamma) \right) \right. \]
\[ + \sin G(z/x^\gamma) \sin G(z/\tilde{x}^\gamma) (v - v_*) \cdot (\tilde{v} - \tilde{v}_*) \]
\[ = \frac{1}{4} \left[ (1 - \cos G(z/x^\gamma)) + (1 - \cos G(z/\tilde{x}^\gamma)) \right. \]
\[ - \left. \left( 1 - \cos \left( G(z/x^\gamma) - G(z/\tilde{x}^\gamma) \right) \right) \right] (v - v_*) \cdot (\tilde{v} - \tilde{v}_*) \tag{211} \]

Similar, elementary calculations show that
\[ |a|^2 = \frac{1}{2} (1 - \cos G(z/x^\gamma)) |v - v_*|^2; \quad |\tilde{a}|^2 = \frac{1}{2} (1 - \cos G(z/\tilde{x}^\gamma)) |\tilde{v} - \tilde{v}_*|^2. \tag{212} \]

We now observe that
\[ \int_0^\infty (1 - \cos G(z/x^\gamma)) dz = cx^\gamma \tag{213} \]
and so, from (211, 212), we obtain
\[ \int_0^\infty dz \int_{S^{d-2}} d\varphi |a - \tilde{a}|^2 \leq \frac{c}{2} \left( x^{2+\gamma} + \tilde{x}^{2+\gamma} - (v - v_*) \cdot (\tilde{v} - \tilde{v}_*) (x^\gamma + \tilde{x}^\gamma) \right. \]
\[ + \frac{x\tilde{x}}{4} \int_0^\infty \left( G(z/x^\gamma) - G(z/\tilde{x}^\gamma) \right)^2 dz. \tag{214} \]
Recalling that \( x^2 = (v - v_*) \cdot (v - v_*) \) and \( \tilde{x}^2 = (\tilde{v} - \tilde{v}_*) \cdot (\tilde{v} - \tilde{v}_*) \), the term in parentheses on the first line rearranges to

\[
\begin{align*}
    x^{2+\gamma} + \tilde{x}^{2+\gamma} - (v - v_*) \cdot (\tilde{v} - \tilde{v}_*) (x^{\gamma} + \tilde{x}^{\gamma}) = (v - v_*) \cdot [(v - v_*) - (\tilde{v} - \tilde{v}_*)] x^{\gamma}, \\
    + (\tilde{v} - \tilde{v}_*) \cdot [(\tilde{v} - \tilde{v}_*) - (v - v_*)] \tilde{x}^{\gamma} = ((v - \tilde{v}) - (v_* - \tilde{v}_*)) \cdot [(v - v_*) x^{\gamma} - (\tilde{v} - \tilde{v}_*) \tilde{x}^{\gamma}].
\end{align*}
\]  
(215)

The same estimates as in (207) now give

\[
x^{2+\gamma} + \tilde{x}^{2+\gamma} - (v - v_*) \cdot (\tilde{v} - \tilde{v}_*) (x^{\gamma} + \tilde{x}^{\gamma}) \leq c \left( |v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2 \right) \left( |v|^{\gamma} + |v_*|^{\gamma} + |\tilde{v}|^{\gamma} + |\tilde{v}_*|^{\gamma} \right).
\]  
(216)

Let us now consider the final line of (214). By Lemma 12, we have the bound

\[
\int_0^\infty \left( G(z/x^{\gamma}) - G(z/\tilde{x}^{\gamma}) \right)^2 \, dz \leq c \frac{|x^{\gamma} - \tilde{x}^{\gamma}|^2}{x^{\gamma} + \tilde{x}^{\gamma}}.
\]  
(217)

We therefore obtain

\[
\begin{align*}
    x\tilde{x} \int_0^\infty \left( G(z/x^{\gamma}) - G(z/\tilde{x}^{\gamma}) \right)^2 \, dz &\leq c \frac{\min(x, \tilde{x})}{\max(x, \tilde{x})^{1-\gamma}} |x - \tilde{x}|^2 \leq c(|v|^{\gamma} + |v_*|^{\gamma} + |\tilde{v}|^{\gamma} + |\tilde{v}_*|^{\gamma})(|v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2).
\end{align*}
\]  
(218)

Combining (214, 216, 218), we have shown that

\[
H_3 \leq c(|v|^{\gamma} + |v_*|^{\gamma} + |\tilde{v}|^{\gamma} + |\tilde{v}_*|^{\gamma})(|v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2) f_p^*(|v|, |v_*|).
\]  
(219)

4. Analysis of \( H_4 \). The final error term is the term \( T_4 \), which corresponds to collisions in the noncutoff system with no corresponding event in the cutoff system. As a result, we anticipate that \( H_4 \) will not be bounded in terms of \( v - \tilde{v}, v_* - \tilde{v}_* \), but will be small in the limit \( K \to \infty \). Let us recall that the integration limit \( L \) is defined as \( L := K \tilde{x}^{\gamma} \). By expanding out the norms, we bound the integrand, for \( z \geq L \),

\[
\begin{align*}
    f_p^* (|v|, |v_*|, z, x) |v + a - \tilde{v}|^2 - |v + a - \tilde{v} - \tilde{\alpha}|^2 &\leq c f_p^* (|v|, |v_*|) |\tilde{\alpha}| (|v| + |\tilde{v}| + |v_*| + |\tilde{v}_*|).
\end{align*}
\]  
(220)

As above, we have

\[
|\tilde{\alpha}| = \sqrt{\frac{1}{2} \left( 1 - \cos G \left( \frac{z}{\tilde{x}^{\gamma}} \right) \right)} |\tilde{v} - \tilde{v}_*| \leq \frac{1}{2} G \left( \frac{z}{\tilde{x}^{\gamma}} \right) |\tilde{v} - \tilde{v}_*|.
\]  
(221)

We therefore obtain the bound

\[
H_4 \leq c f_p^* (|v|, |v_*|) (|v|^2 + |v_*|^2 + |\tilde{v}|^2 + |\tilde{v}_*|^2) \int_L^\infty G \left( \frac{z}{\tilde{x}^{\gamma}} \right) \, dz.
\]  
(222)

Recalling the definition of \( L = K \tilde{x}^{\gamma} \), the integral evaluates to

\[
\int_L^\infty G \left( \frac{z}{\tilde{x}^{\gamma}} \right) \, dz = \tilde{x}^{\gamma} \int_K^\infty G(z) \, dz \leq c \tilde{x}^{\gamma} K^{1-1/\gamma}.
\]  
(223)
We therefore find
\[ H_4 \leq c K^{1 - 1/\nu} (|v|^{p+2+\gamma} + |\tilde{v}|^{p+2+\gamma} + |v_*|^{p+2+\gamma} + |\tilde{v}_*|^{p+2+\gamma}) \] (224)

Recalling that \( l := p + 2 + \gamma \), this is exactly the error claimed.

5. Converting into the form desired. Combining (200, 209, 219, 224), we see that
\[
\mathcal{E}_{p,K}^1 \leq \left( c - \frac{\lambda p}{2} \right) |v|^{p+\gamma} + C |v_*|^{p+\gamma} \right) |v - \tilde{v}|^2 + c(|v|^{\gamma} + |\tilde{v}|^{\gamma} + |v_*|^{\gamma} \\
+ |\tilde{v}_*|^{\gamma}) f_p^*(|v|, |v_*|)(|v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2) \\
+ c(|v|^{l} + |\tilde{v}|^{l} + |v_*|^{l} + |\tilde{v}_*|^{l}) K^{1-1/\nu}. \] (225)

The first and last lines are already in the form desired in the statement of the lemma. For the middle line, we use Peter-Paul, noting that different bounds are required depending on whether we multiply \(|v - \tilde{v}|^2\) or \(|v_* - \tilde{v}_*|^2\); it is essential that the coefficients multiplying \(|v|^{p+\gamma}|v - \tilde{v}|^2\) and \(|v_*|^{p+\gamma}|v_* - \tilde{v}_*|^2\) do not depend on \( p \). We start from \( f_p^*(|v|, |v_*|) \leq 2 |v|^p + C |v_*|^p \), and note also from Young’s inequality that, for all \( a, b > 0 \), \( a b^\nu \leq a^{p+\gamma} + b^{p+\gamma} \). Therefore,
\[
(|v|^\gamma + |v_*|^\gamma + |\tilde{v}|^\gamma + |\tilde{v}_*|^{\gamma}) f_p^*(|v|, |v_*|) \\
\leq c(|v|^{p+\gamma} + |v_*|^{p+\gamma} + |\tilde{v}|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma}) \\
+ C(|v_*|^p(|v|^\gamma + |\tilde{v}|^{\gamma}) + |v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma}) \\
\leq c |v|^{p+\gamma} + c |\tilde{v}|^{p+\gamma} + C(1 + |v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma})(1 + |v|^p + |\tilde{v}|^p). \] (226)

We use this inequality for the term multiplying \(|v - \tilde{v}|^2\), and reverse the roles of \( v \leftrightarrow v_* \), \( \tilde{v} \leftrightarrow \tilde{v}_* \) for the term involving \(|v_* - \tilde{v}_*|^2\). Together, we see that
\[
(|v|^\gamma + |\tilde{v}|^\gamma + |v_*|^\gamma + |\tilde{v}_*|^{\gamma}) f_p^*(|v|, |v_*|)(|v - \tilde{v}|^2 + |v_* - \tilde{v}_*|^2) \\
\leq c(|v|^{p+\gamma} + |\tilde{v}|^{p+\gamma}) |v - \tilde{v}|^2 + c(|v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma}) |v_* - \tilde{v}_*|^2 \\
+ C(1 + |v_*|^{p+\gamma} + |\tilde{v}_*|^{p+\gamma})(1 + |v|^p + |\tilde{v}|^p)|v - \tilde{v}|^2 \\
+ C(1 + |v|^p + |\tilde{v}|^p + |\tilde{v}_*|^{p+\gamma})(1 + |v_*|^p + |\tilde{v}_*|^{p+\gamma}) |v_* - \tilde{v}_*|^2 \] (227)

which gives the bound desired for \( \mathcal{E}_{p,K}^1 \).

6. Estimate on \( \mathcal{E}_{p,K}^2 \). We now turn to the analysis of \( \mathcal{E}_{p,K}^2 \), which follows a similar pattern to \( \mathcal{E}_{p,K}^1 \) above. In this case, we use the bound
\[
|\tilde{v} + \tilde{a}_K|^p \leq f_{p,L}(|\tilde{v}|, |\tilde{v}_*|, z, \tilde{x}) = f_p(|\tilde{v}|, |\tilde{v}_*|, z, \tilde{x}), \quad z \leq L; \\
|\tilde{v}|^p, \quad z > L \] (228)

which has the same upper bound \( f_p^* \). We therefore obtain a decomposition equivalent to (193):
\[
\mathcal{E}_{p,K}^2 \leq \tilde{H}_1 + \tilde{H}_2 + \tilde{H}_3 + \tilde{H}_4 \] (229)
where
\[
\begin{align*}
\tilde{H}_1 & := \int_0^\infty dz \left( f_{p,L}(|\tilde{v}|, |\tilde{v}_*|, z, x) - |\tilde{v}|^p \right) |v - \tilde{v}|^2; \\
\tilde{H}_2 & := 2 \int_0^L dz \int_{S_{d-2}} d\varphi \: f_{p,L}(|\tilde{v}|, |\tilde{v}_*|, z, x) (v - \tilde{v}) \cdot (a - \tilde{a}); \\
\tilde{H}_3 & := \int_0^\infty dz \int_{S_{d-2}} d\varphi \: f_p(|\tilde{v}|, |\tilde{v}_*|)(a - \tilde{a})^2; \\
\tilde{H}_4 & := \int_L^\infty dz \int_{S_{d-2}} d\varphi \: f_{p,L}(|\tilde{v}|, |\tilde{v}_*|, z, \tilde{x}) \left| 2(v - \tilde{v}) \cdot \tilde{a} + |\tilde{a}|^2 \right|
\end{align*}
\] (230)

The analyses of \( \tilde{H}_3, \tilde{H}_4 \) are identical to the arguments above, and we will now discuss the necessary modifications for \( \tilde{H}_1, \tilde{H}_2 \).

\(6a.\) Analysis of \( \tilde{H}_1 \). Let us begin with \( \tilde{H}_1 \). The same reparametrisation gives
\[
\begin{align*}
\int_0^\infty (f_{p,L}(|\tilde{v}|, |\tilde{v}_*|, z, \tilde{x}) - |\tilde{v}|^p) dz \\
& \leq -|\tilde{v} - \tilde{v}_*|^\gamma |\tilde{v}|^p \int_{\theta_0(K)}^{\pi/2} \left( 1 - \left( \frac{1 + \cos \theta}{2} \right)^{p/2} \right) \beta(\theta) d\theta + |\tilde{v} - \tilde{v}_*|^\gamma (|\tilde{v}_*|^p + C(|\tilde{v}|^{p-2}|\tilde{v}_*|^2 + |\tilde{v}|^2|\tilde{v}_*|^{p-2})).
\end{align*}
\] (234)

We therefore obtain
\[
\begin{align*}
\int_0^\infty (f_{p,L}(|\tilde{v}|, |\tilde{v}_*|, z, \tilde{x}) - |\tilde{v}|^p) dz \\
& \leq -|\tilde{v} - \tilde{v}_*|^\gamma |\tilde{v}|^p \lambda_{p,K} + |\tilde{v} - \tilde{v}_*|^\gamma (|\tilde{v}_*|^p + C(|\tilde{v}|^{p-2}|\tilde{v}_*|^2 + |\tilde{v}|^2|\tilde{v}_*|^{p-2})) \\
& \leq -\lambda_{p,K} |\tilde{v}|^{p+\gamma} + \lambda_p |\tilde{v}_*|^\gamma |\tilde{v}|^p + |\tilde{v} - \tilde{v}_*|^\gamma (|\tilde{v}_*|^p + C(|\tilde{v}|^{p-2}|\tilde{v}_*|^2 + |\tilde{v}|^2|\tilde{v}_*|^{p-2}))
\end{align*}
\] (235)

where
\[
\lambda_{p,K} := \int_{\theta_0(K)}^{\pi/2} \left( 1 - \left( \frac{1 + \cos \theta}{2} \right)^{p/2} \right) \beta(\theta) d\theta \leq \lambda_p.
\] (236)

We now use Peter-Paul, independently of \( K \), to obtain
\[
\lambda_p |\tilde{v}_*|^\gamma |\tilde{v}|^p + |\tilde{v} - \tilde{v}_*|^\gamma (|\tilde{v}_*|^p + C(|\tilde{v}|^{p-2}|\tilde{v}_*|^2 + |\tilde{v}|^2|\tilde{v}_*|^{p-2})) \\
\leq \frac{\lambda_p}{3} |\tilde{v}|^{p+\gamma} + C|\tilde{v}_*|^{p+\gamma}.
\] (237)

By monotone convergence, \( \lambda_{p,K} \to \lambda_p \) as \( K \to \infty \) with \( p \) fixed and, in particular, for \( K \geq K_0(p) \) large enough, \( \lambda_{p,K} \geq \frac{5}{6} \lambda_p \). For such \( K \), we have shown that
\[
\tilde{H}_1 \leq -\frac{\lambda_p}{2} |\tilde{v}|^{p+\gamma} |v - \tilde{v}|^2 + C|\tilde{v}_*|^{p+\gamma} |v - \tilde{v}|^2.
\] (238)
6b. Analysis of $\tilde{H}_2$. Following the same manipulations as (202), we obtain

$$\tilde{H}_2 = (v - \tilde{v}) \cdot \left\{ (\Psi_{0L} + \Psi_{L\infty})(\tilde{x}'', |\tilde{v}|, |\tilde{v}_a|, \tilde{x})(\tilde{v} - \tilde{v}_a) - (\Psi_{0L} + \Psi_{L\infty})(\tilde{x}'', |\tilde{v}|, |\tilde{v}_a|, \tilde{x})(v - v_a) \right\}$$

(239)

where we define

$$\Psi_{0L}(y, u, v, w) = \int_0^L f_p(u, v, z, w)(1 - \cos G(z/y))dz = y \int_0^{L/y} f_p(u, v, zy, w)(1 - \cos G(z))dz$$

(240)

and

$$\Psi_{L\infty}(y, u, v, w) = \int_{L}^{\infty} v^p(1 - \cos G(z/y))dz = y \int_{L/y}^{\infty} v^p(1 - \cos G(z))dz.$$

(241)

One then repeats the differentiation (205); in the case of $\Psi_{0L}$, a surface term appears, and is subsequently removed by integration by parts, while in $\Psi_{L\infty}$ one bounds the surface term by

$$\frac{L}{y} v^p \left( 1 - \cos \left( \frac{L}{y} \right) \right) \leq v^p \sup_{z \geq 0} [z(1 - \cos G(z))]$$

(242)

and this supremum is readily seen to be finite using Lemma 12. Together, one concludes a bound

$$\left| \frac{\partial}{\partial y} \Psi_{0L}(y, u, v, w) \right| + \left| \frac{\partial}{\partial y} \Psi_{L\infty}(y, u, v, w) \right| \leq cf_p^*(u, v)$$

(243)

and the rest of the argument follows as for $H_2$.

7. Bound on $E_{3p,K}$. Finally, let us mention $E_{3p,K}$. This term is strictly easier than the two above; there is no term analogous to $H_1$, and one can omit the moment prefactors in the remaining terms. Alternatively, one may note that $E_{3p,K}$ is exactly that analysed in [20, Lemma 3.1], and the claimed bound is exactly the content of [20, Lemma 5.1].

8.3. Proof of Lemma 6

We now turn to the proof of the quadratic bound Lemma 6, where we replace the integrand of $E_{p,K}$ with its square. In this case, the integrand is nonnegative, and there is no hope of exploiting cancellations in the way we did above. On the other hand, the statement we seek to prove is much weaker; we ask only for local boundedness of $Q_K$, rather than being small in a suitable sense when $|v - \tilde{v}|, |v_a - \tilde{v}_a|$ are small. It will be sufficient to prove the following slightly simpler lemma, which breaks up $Q_K$ in a similar way to the decomposition $E_{p,K} = E_{1p,K} + E_{2p,K} + E_{3p,K}$ above:
Lemma 14. Define

\[ Q^1_K = \int_0^\infty dz \int_{S^{d-2}} d\varphi \left( \frac{d^2}{p} (v', \tilde{v}'_K) - \frac{d^2}{p} (v, \tilde{v}) \right)^2; \]  
\[ Q^2_K = \int_0^\infty dz \int_{S^{d-2}} d\varphi \left( \frac{d^2}{p} (v'_*, \tilde{v}'_K) - \frac{d^2}{p} (v_*, \tilde{v}_*) \right)^2. \]  

Then the estimate (82) holds with either \( Q^1_K \) or \( Q^2_K \) in place of \( Q_K \).

Once we have established these estimates, the second point of Lemma 6 follows from the easy comparison \( Q^2_K \leq 2Q^2_{K,1} + 2Q^2_{K,1} \).

Proof of Lemmata 6. We use the same notation as above, and start from a decomposition similar to (193):

\[ d^2_p (v', \tilde{v}'_K) - d^2_p (v, \tilde{v}) = (|v'|^p + |\tilde{v}'_K|^p - |v|^p - |\tilde{v}|^p) |v - \tilde{v}|^2 + (1 + |v'|^p + |\tilde{v}'_K|^p) (2(a - \tilde{a}) \cdot (v - \tilde{v}) + |a - \tilde{a}|^2) \]
\[ + (1 + |v'|^p + |\tilde{v}'_K|^p) (2\tilde{a} \cdot (v + a - \tilde{v}) + |\tilde{a}|^2) 1(z \geq L). \]

We now square each term, and use the crude bounds \( |a| \leq |v| + |v_*|, |\tilde{a}| \leq |\tilde{v}| + |\tilde{v}_*| \) to see that

\[ (d^2_p (v', \tilde{v}'_K) - d^2_p (v, \tilde{v}))^2 \leq c(|v'|^p + |\tilde{v}'_K|^p - |v|^p - |\tilde{v}|^p)^2 |v - \tilde{v}|^4 + c(1 + |v'|^p + |\tilde{v}'_K|^p)^2 (|v|^2 + |\tilde{v}|^2 + |v_*|^2 + |\tilde{v}_*|^2) |a - \tilde{a}|^2 \]
\[ + c(1 + |v'|^p + |\tilde{v}'_K|^p)^2 (|v|^2 + |\tilde{v}|^2 + |v_*|^2 + |\tilde{v}_*|^2) |\tilde{a}|^2 1(z \geq L). \]

We can now replace every instance of \( |v'|^p \leq C(|v|^p + |v_*|^p) \), and similarly for \( \tilde{v}'_K \), and drop the factor \( 1(z \geq L) \) in the final term. In this way, we obtain

\[ Q^1_K \leq C (H_5 + H_6 + H_7); \]

where the three terms are

\[ H_5 := \int_0^\infty dz \int_{S^{d-2}} d\varphi \left| |v'|^p - |v|^p + |\tilde{v}'_K|^p - |\tilde{v}|^p \right| \]
\[ \cdots \times (1 + |v|^p + |v_*|^p + |\tilde{v}| + |\tilde{v}_*| + |\tilde{a}|); \]
\[ H_6 := \int_0^\infty dz \int_{S^{d-2}} d\varphi (1 + |v|^{2p+2} + |v_*|^{2p+2} + |\tilde{v}|^{2p+2} + |\tilde{v}_*|^{2p+2}) |a - \tilde{a}|^2; \]
\[ H_7 := \int_0^\infty dz \int_{S^{d-2}} d\varphi (1 + |v|^{2p+2} + |v_*|^{2p+2} + |\tilde{v}|^{2p+2} + |\tilde{v}_*|^{2p+2}) |\tilde{a}|^2. \]

Let us now analyse these integrals one by one. The analysis of \( H_5 \) is similar to that of \( H_1 \), although with an absolute value, and the integrals appearing in \( H_6, H_7 \) can
be reduced to the calculations for \( H_3, H_4 \) in the previous proof. 1. Analysis of \( H_5 \). We start from the observation that, for all \( v, w \in \mathbb{R}^d \), we have
\[
||v|^p - |w|^p| \leq C(1 + |v|^{p-1} + |w|^{p-1})|v - w|.
\] (252)

It follows that
\[
||v'|^p - |v|^p| \leq C(1 + |v|^{p-1} + |v + a|^{p-1})|a|
\leq C(1 + |v|^{p-1} + |v_*|^{p-1})(|v| + |v_*|)G\left(\frac{z}{\chi^\gamma}\right).
\] (253)

Integrating, we find that
\[
\int_0^\infty dz \int_{\mathbb{S}^{d-2}} d\varphi \, ||v'|^p - |v|^p| \leq C(1 + |v|^p + |v_*|^p) \int_0^\infty G(z/\chi^\gamma)dz
\leq C(1 + |v|^{p+\gamma} + |v_*|^{p+\gamma}).
\] (254)

A similar argument applies for \( ||\tilde{v}'_v|^p - |	ilde{v}|^p| \). Including the moment prefactors, we obtain
\[
H_5 \leq C(1 + |v|^{2p+4+\gamma} + |v_*|^{2p+4+\gamma} + |\tilde{v}|^{2p+4+\gamma} + |\tilde{v}_*|^{2p+4+\gamma}).
\] (255)

2. Analysis of \( H_6 \). For \( H_6 \), we note that the moment prefactor is constant over the integral, and that we already analysed \( \int_0^\infty dz \int_{\mathbb{S}^{d-2}} |a - \tilde{a}|^2 \) when analysing \( H_3 \) in the previous proof. Absorbing the terms \( |v - \tilde{v}|^2 \) and \( |v_* - \tilde{v}_*|^2 \), the same calculations as above therefore give
\[
Q_2 \leq C(1 + |v|^{2p+4+\gamma} + |v_*|^{2p+4+\gamma} + |\tilde{v}|^{2p+4+\gamma} + |\tilde{v}_*|^{2p+4+\gamma}).
\] (256)

3. Analysis of \( H_7 \). As above, the moment prefactor is independent of the integration variables \( z, \varphi \), and the problem reduces to estimating \( \int_L^\infty \int_{\mathbb{S}^{d-2}} |\tilde{a}|^2 \), which is analagous to \( H_4 \). We recall that
\[
|\tilde{a}|^2 = \frac{1}{2}|\tilde{v} - \tilde{v}_*|^2 \left(1 - \cos G\left(\frac{z}{\chi}\right)\right) \leq \frac{1}{4}|v - \tilde{v}|^2 G\left(\frac{z}{\chi^\gamma}\right)^2.
\] (257)

Therefore,
\[
\int_0^\infty dz \int_{\mathbb{S}^{d-2}} d\varphi |\tilde{a}|^2 \leq C|v - \tilde{v}|^2|v - \tilde{v}|^\gamma \int_0^\infty G(z)dz.
\] (258)

The final integral is finite, thanks to the estimates established in Section 8.1, so we conclude
\[
H_7 \leq C(1 + |v|^{2p+4+\gamma} + |v_*|^{2p+4+\gamma} + |\tilde{v}|^{2p+4+\gamma} + |\tilde{v}_*|^{2p+4+\gamma}).
\] (259)

Combining (255, 256, 259) gives the claimed result. \qed

Acknowledgements. I am very grateful to my doctoral supervisor, James Norris, for the suggestion of this project and for several useful conversations. This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) Grant EP/L016516/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis.
A. Appendix: Convergence and Stability Estimates for the Cutoff Kac Process and Boltzmann Equation

We now give the proof of Lemmata 2, 3, which we deferred earlier. The strategy for the whole proofs follows Norris [34], with Lemma 2 corresponding to [34, Theorem 1.1], and Lemma 3 corresponding to [34, Theorem 10.1], and so very few points are novel. In order to discuss points which differ from Norris’ treatment, and to keep the present work as self-contained as possible, we will give statements of the intermediate steps and indicate to which results in the cited work these statements correspond. The details of these arguments may be found in the author’s doctoral thesis [25, Section 3.2.1–3.2.2, pp. 112–128].

We first define a metric which is well-suited to the duality technique in question. For \( f : \mathbb{R}^d \to \mathbb{R} \), define a rescaled function \( \hat{f}(v) = f(v)/(1 + |v|^2) \), and the \( \gamma \)-Hölder norm

\[
\| f \|_{0, \gamma} := \max \left( \sup_v |f(v)|, \sup_{v \neq w} \frac{|f(v) - f(w)|}{|v - w|^{\gamma}} \right).
\]  

We write \( \mathcal{A}_\gamma \) for the space of weighted \( \gamma \)-Hölder functions where this norm is at most 1:

\[
\mathcal{A}_\gamma := \left\{ f : \mathbb{R}^d \to \mathbb{R} : \| \hat{f} \|_{0, \gamma} \leq 1 \right\}
\]

and define the weighted Wasserstein metric of type \( \gamma \) by the duality

\[
\mathcal{W}_\gamma(\mu, \nu) := \sup_{f \in \mathcal{A}_\gamma} |\langle f, \mu - \nu \rangle|.
\]

A chain of elementary estimates [25, Section 2.1.4] shows that, if \( p \geq 0 \), \( p' > p + 2 \) and \( 0 < \gamma \leq 1 \), we have the equivalence

\[
\mathcal{W}_\gamma(\mu, \nu) \leq C \Lambda_{p'}(\mu, \nu) T_p(\mu, \nu)^{\alpha}; \quad T^2_p(\mu, \nu) \leq C \Lambda_p(\mu, \nu) \mathcal{W}_\gamma(\mu, \nu)^{\alpha}
\]

for some \( C < \infty \), \( \alpha > 0 \) depending on \( p \), \( p' \), \( \gamma \).

A.1. Random measures associated to the cutoff process

We begin by first introducing the jump measure and compensator associated to the cutoff Kac process \( (\mu^N_t, K_t)_{t \geq 0} \).

Definition 1. (Jump Measure and Compensator) Let \( (\mu^N_t, K_t)_{t \geq 0} \) be a cutoff Kac process on \( N \) particles.
(i) The jump measure $m^{N,K}_\infty$ is the unnormalised empirical measure on $(0, \infty) \times S_N$ on the set of pairs $(t, \mu_{t^+}^{N,K})$ such that $\mu_{t}^{N,K} \neq \mu_{t}^{N,K}_{-}$.

(ii) Let $Q_{N,K}$ be the kernel on $S_N$ given by

$$Q_{N,K}(\mu^N, A) = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}^{d-1}} 1(\mu^N(v, v^*, \sigma) \in A) B_K(v - v^*, d\sigma) \mu^N(dv) \mu^N(dv^*).$$

The compensator $m^{N,K}$ of the jump measure is the measure on $(0, \infty) \times S_N$ given by

$$m^{N,K}(dt, A) = Q_{N,K}(\mu^N, A) dt.$$  

Since we are working with the cutoff process, both of these measures are almost surely finite on compact subsets $(0, t) \times S_N$, for any $t < \infty$.

### A.2. The linearised kac process for the cutoff case

We next sketch the proof of Lemma 3. In order to apply the ideas of [34], we start from the following continuity property of the kernel.

**Lemma 15.** There exists $C < \infty$, $\alpha > 0$, depending only on $b$ such that, for all $v, w \in \mathbb{R}^d$ and all $K \geq 1$, we have the estimate

$$\sup_{v^* \in \mathbb{R}^d} \|B_K(v - v^*, \cdot) - B_K(w - v^*, \cdot)\|_{L^1(d\sigma)} \leq CK^\alpha |v - v'|.$$  

We refer to [25, Lemma 3.12, p. 116] for a proof. Following the ideas of [34, Section 4], we introduce the following linearised Kac process. Throughout, we fix $K \geq 1$, and omit it from the notation.

**Definition 2.** Let us write $V = \mathbb{R}^d$ and $V^*$ for the signed space $V^* = V \times \{\pm 1\} = V^+ \cup V^-$. We write $\pi : V^* \to V$ as the projection onto the first factor, and $\pi_\pm : V^\pm \to V$ for the obvious bijections. Let also $\rho = (\rho_t)_{t \geq 0}$ be family of measures on $V = \mathbb{R}^d$ such that

$$\langle 1, \rho_t \rangle = 1; \quad \langle |v|^2, \rho_t \rangle = 1 \quad \text{for all } t \geq 0;$$

$$\int_0^t \Lambda_{2+\gamma}(\rho_s) ds < \infty \quad \text{for all } t < \infty.$$  

The Linearised Kac Process in environment $(\rho_t)_{t \geq 0}$ is the branching process on $V^*$ where each particle of type $(v, 1)$, at rate $2B_K(v - v^*, d\sigma) \rho_t(dv^*)$, dies, and is replaced by three particles, of types

$$(v^*(v, v^*, \sigma), 1); \quad (v^*(v, v^*, \sigma), 1); \quad (v^*, -1)$$

where $v', v^*$ are the post-collisional velocities. The dynamics are identical for particles of type $(v, -1)$, with the signs exchanged.

We write $E_t^*$ for the associated process of unnormalised empirical measures on $V^*$, and define a signed measure $E_t$ on $V$ by including the sign at each particle:

$$E_t = E_t^+ - E_t^-; \quad E_t^\pm = E_t^* \circ \pi_\pm^{-1}.$$  

We can also consider the same branching process, started from a time $s \geq 0$ instead. We write $E$ for the expectation over the branching process, which is not the full expectation in the case where $\rho$ is itself random. When we wish to emphasise the initial velocity $v$ and starting time $s$, we will write $E_{(s,v)}$ when the process is started from $E_0^* = \delta_{(v,1)}$ at time $s$, and $E_v$ in the case $s = 0$. 

Provided that the initial data $Ξ^*_0$ satisfies $E(1 + |v|^2, |Ξ_0|) < \infty$, one can show that the branching process $(Ξ_t)_{t \geq 0}$ is non-explosive, and therefore defined for all time $t \geq 0$. Moreover, the bound is propagated:

$$E(1 + |v|^2, Ξ^*_t) \leq \exp \left( CK \int_0^t \Lambda_{2+\gamma}(\rho_s) ds \right) E(1 + |v|^2, Ξ^*_0).$$  \hspace{1cm} (271)

We can therefore define, for functions $f$ of quadratic growth,

$$f_{st}(v) = E_{(s,v)}(f, Ξ_t).$$  \hspace{1cm} (272)

We will write $f_{st}[\rho](v)$ when we wish to emphasise the dependence on the environment $\rho$. The proof of Lemma 3 is based on the following representation formula, which can be proved with only slight modifications of [34, Proposition 4.2].

**Proposition 8.** (Representation formula for Cutoff Cases) Let us fix $\mu, \nu \in S^{2+\gamma}$, and consider the environment

$$\rho_t = \phi_{Kt}(\mu) + \phi_{Kt}(\nu)^2.$$  \hspace{1cm} (273)

Then, for all $t \geq 0$ and all $f$ of quadratic growth,

$$\langle f, \phi_{Kt}^i(\mu) - \phi_{Kt}^i(v) \rangle = \langle f_{0t}[\rho], \mu - \nu \rangle.$$  \hspace{1cm} (274)

Further, let $\mu_{N,k}^N$ be a cutoff Kac process on $N$ particles, and let $m_{N,k}^N, \bar{m}_{N,k}^N$ be its jump measure and compensator, as in Definition 1. In this case, consider propagation $f_{st} = f_{st}[\rho^N]$ in the random environment

$$\rho_t^N = \mu_{t,k}^N + \phi_{Kt}(\mu)^2.$$  \hspace{1cm} (275)

Then, for all $t \geq 0$, and all functions $f$ of quadratic growth, we have

$$\langle f, \mu_{t,k}^N - \phi_{Kt}^i(\mu) \rangle = \langle f_{0t}[\rho^N], \mu_{t,k}^N - \mu \rangle + M_{t,k}^N.f$$  \hspace{1cm} (276)

where

$$M_{t,k}^N.f = \int_{[0,t] \times S_N} \langle f_{st}[\rho^N], \mu^N - \mu_{s,k}^N \rangle (m_{N,k}^N - \bar{m}_{N,k}^N)(ds, d\mu^N).$$  \hspace{1cm} (277)

This is essentially identical to [34, Proposition 4.2], see also [25, Proof of Proposition 3.10, p. 119].

Recalling the definition of $w_{\gamma}$ in (261,262), the proof of Lemma 3 will be to estimate growth and regularity of $f_{0t}$ when we start with a function $f \in A_{\gamma}$, which leads to stability estimates in $w_{\gamma}$. The following result adapts [34, Proposition 4.3] to our case.

**Lemma 16.** (Growth and Regularity of $f_{0t}$) Fix $f \in A_{\gamma}$ and an environment $\rho_t, t \geq 0$. Then $f_{0t} \in z_t A_{\gamma}$, where

$$z_t = \exp \left( CK \left( 1 + \int_0^t \Lambda_{2+\gamma}(\rho_s) ds \right) \right)$$  \hspace{1cm} (278)

for some constant $C$ independent of $K$.

This lemma may be proven by repeating the calculations leading to [34, Propositions 4.3, 4.5], and keeping track of the dependence on $K$. For the estimate of the continuity, one replaces [34, eq. (2)] with the estimate Lemma 15 presented above. Details of the necessary modifications for the case considered here can be found in [25, Proposition 3.9, pp. 115, 117–119].
A.3. Proof of Lemmata 2, 3

Combining the previous lemmata, we prove the two Lemmata 2, 3. The estimates given for the branching process functions \( f_{st} \) above, and the representation lemma, naturally lead to estimates in the weighted metric \( w_\gamma \) given by optimising over \( f \) in a given weighted-Hölder class; this is the argument of Norris, who works in (in our notation) \( w_1 \). To convert everything into the statements measured in \( T_\rho \) given in the text, we will use the interpolation estimates in (263).

We begin with Lemma 3, which is a simple application of Proposition 8 and the estimates in Lemma 16, and using the interpolations between the and our optimal transport cost \( T_\rho \).

**Proof of Lemma 3.** Let us fix \( \mu, v \in S^p \), for \( p \) to be chosen later. Let \( \rho \) be the environment

\[
\rho_t = \frac{1}{2} (\phi^K_t(\mu) + \phi^K_t(v)) \tag{279}
\]

and let \( f_{st} \) denote the functions given by (272) in this environment. For any \( f \in A_\gamma \), we have

\[
\langle f, \phi^K_t(\mu) - \phi^K_t(v) \rangle = \langle f_{0t}, \mu - v \rangle \leq z_t \ w_\gamma(\mu, v) \tag{280}
\]

where \( z_t \) is as in Lemma 16; by Proposition 4, we bound

\[
z_t \leq \exp \left( cK(1 + t) \Lambda_{2+\gamma}(\mu, v) \right) \tag{281}
\]

and so, optimising over \( f \),

\[
w_\gamma(\phi^K_t(\mu), \phi^K_t(v)) \leq \exp \left( cK(1 + t) \Lambda_{2+\gamma}(\mu, v) \right) w_\gamma(\mu, v). \tag{282}
\]

Finally, we use (263) twice to convert both sides from \( w_\gamma \) to \( T_\rho \): for some \( C = C(G, p, q, d) \), \( \beta = \beta(p, p') \), \( \beta' = \beta'(p', p') \),

\[
T_\rho \left( \phi^K_t(\mu), \phi^K_t(v) \right) \leq C \Lambda_{p'}(\mu, v) \ w_\gamma \left( \phi^K_t(\mu), \phi^K_t(v) \right) \ w_\gamma(\mu, v)^{\beta'} \leq C \Lambda_{p'}(\mu, v) \ w_\gamma(\mu, v)^{\beta'} \exp \left( cK(1 + t) \Lambda_{2+\gamma}(\mu, v) \right) T_\rho(\mu, v)^{\beta'} \tag{283}
\]

which proves the claim for a new choice of \( \beta \). \( \square \)

For the case with an \( N \)-particle Kac process, we will need to control the stochastic integral term, uniformly over \( f \) belonging to the class of test functions \( A_\gamma \) in the definition (261,262) of \( w_\gamma \). This is achieved with the following proposition.

**Proposition 9.** Let \( \mu_{t}^{N,K}, t \geq 0 \) be a \( N \)-particle, \( K \)-cutoff Kac process, and let \( \mathcal{F}_{t}^{N} \) be the natural filtration. Let \( \rho = (\rho_t)_{t \geq 0} \) be a potentially random environment, adapted to \( \mathcal{F}_{t}^{N} \), such that

\[
\lambda^* = \sup_{t \geq 0} \Lambda_{2+\gamma}(\rho_t) \|_{L^\infty(\mathbb{F})} < \infty. \tag{284}
\]

For \( f \in A_\gamma \) and \( t \geq s \geq 0 \), let \( f_{st}[\rho] \) denote the propagation in this environment, as described in Definition 2. Let \( q \geq 2 + \gamma \) and \( a \geq 1 \), and suppose that \( \mu_{t}^{N,K} \) has an initial moment \( \Lambda_q(\mu_{0}^{N,K}) \leq a \). Let \( m^{N,K}, \overline{m}^{N,K} \) be as in Definition 1, and write

\[
\tilde{M}_{t}^{N,K,f}[\rho] = \int_{[0,t] \times S_N} \langle f_{st}[\rho], \mu^{N} - \mu_{s}^{N,K} \rangle (m^{N,K} - \overline{m}^{N,K}) (ds, d\mu^{N}). \tag{285}
\]
In this notation, we have the bound
\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} \sup_{f \in A_{\gamma}} \tilde{M}^{N,K,f}_{t} \right] \leq C a N^{-\eta} \exp \left( C \lambda^{*} K (1 + t_{\text{fin}}) \right)
\]
(286)
for some \( C = C(d, p, \beta) \) and \( \eta = \eta(d, p) > 0 \). Here, we emphasise that \( \mathbb{E} \) refers to the simultaneous expectation with simultaneous over \( \mu^{N,K}_{t} \) and the environment \( \rho \).

This largely follows the same arguments as the proof of [34, Theorem 1.1, pp. 19–22], and is spelled out in detail in [25, Proposition 3.11, pp. 122–126]. The key difference is that the hypotheses on the environment \( \rho \) guarantee an \( L^{\infty}(\mathbb{P}) \) control on the quantities
\[
z_{t} = \exp \left( c K \int_{0}^{t} \Lambda_{2+\gamma} (\rho_{u}) \, du \right) \quad (287)
\]
and
\[
y_{\beta}(t) = z_{t} \sup_{0 \leq s \leq s' \leq t} \left[ (s' - s)^{-1} \int_{s}^{s'} \Lambda_{2+\gamma} (\rho_{u}) \, du \right], \quad 0 < \beta \leq 1 \quad (288)
\]
in terms of the multiplicative factor \( e^{C \lambda^{*} K (1+t)} \), which describe the continuity of \( f_{s,t} [\rho](v) \) in \( v \) and \( s \) respectively. Finally, we indicate how these results may be used to prove Lemma 2.

**Sketch Proof of Lemma 2.** Let us consider the linearised Kac process in the random environment
\[
\rho^{N}_{t} = \frac{\mu^{N,K}_{t} + \phi^{K}_{t} (\mu^{N,K}_{0})}{2}
\]
(289)
as in Lemma 8, and for \( b \geq 1 \), consider the stopping times \( T_{b}^{N} \) defined in (50) for the \((2+\gamma)\)th moment. Let us write \( \tilde{M}^{N,K,f,b}_{t} \) for the stochastic integrals in (285) in the environment \( \rho^{T_{b}^{N}} \).

We consider the events \( \{T_{b} \leq t_{\text{fin}}\}, \{T_{b} > t_{\text{fin}}\} \) separately. On the event \( \{T_{b} > t_{\text{fin}}\} \), we have the equalities
\[
M^{N,K,f}_{t} = \tilde{M}^{N,K,f,b}_{t}
\]
(290)
while on \( \{T_{b} \leq t_{\text{fin}}\} \) we have the trivial bound
\[
\sup_{t \leq t_{\text{fin}}} w_{\gamma} (\mu^{N,K}_{t}, \phi^{K}_{t} (\mu)) \leq 4.
\]
(291)
Combining, we have the bound
\[
\sup_{t \leq t_{\text{fin}}} w_{\gamma} \left( \mu^{N,K}_{t}, \phi^{K}_{t} (\mu^{N,K}_{0}) \right) \leq \sup_{f \in A_{\gamma}, t \leq t_{\text{fin}}} \left\{ \tilde{M}^{N,K,f,b}_{t} \right\} + 4 \cdot 1(T_{b}^{N} \leq t_{\text{fin}}).
\]
(292)
Since \( q > 2 + \gamma \), the moment hypothesis on \( \mu^{N,K}_{0} \) implies \( \Lambda_{2+\gamma} (\mu^{N,K}_{0}) \leq a \) almost surely, which is propagated to \( \phi^{K}_{t} (\mu_{0}) \) by Proposition 4. The first term is therefore controlled by Proposition 9, with \( \lambda^{*} \leq b + C a \) for some constant \( C \). We now take \( b = C a \), for some large \( C \); by Lemma 1, \( C \) can be chosen so that \( \mathbb{P}(T_{b}^{N} \leq t_{\text{fin}}) \leq C a N^{-1} t_{\text{fin}} \). Combining, we obtain
\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} w_{\gamma} \left( \mu^{N,K}_{t}, \phi^{K}_{t} (\mu^{N,K}_{0}) \right) \right] \leq C a N^{-\eta} \exp(Ca K (1 + t_{\text{fin}})) + Cat_{\text{fin}} N^{-1}
\]
(293)
and keeping the worse term
\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} \mathcal{W}_Y \left( \mu_{t}^{N,K}, \phi_t^K(\mu_0^{N,K}) \right) \right] \leq C a N^{-\eta} \exp(C a K (1 + t_{\text{fin}})).
\]  
(294)

To convert this approximation into \( T_p \), we argue as in (283). Fix \( p' \in (p + 2, q) \); thanks to the comparisons in (263), for some \( \alpha > 0 \),
\[
\sup_{t \leq t_{\text{fin}}} T_p \left( \mu_{t}^{N,K}, \phi_t^K(\mu_0^{N,K}) \right) \leq \left( \sup_{t \leq t_{\text{fin}}} \mathcal{W}_Y \left( \mu_{t}^{N,K}, \phi_t^K(\mu_0^{N,K}) \right) \right)^{\alpha} \left( \sup_{t \leq t_{\text{fin}}} \Lambda_{p'} \left( \mu_{t}^{N,K}, \phi_t^K(\mu_0^{N,K}) \right) \right). 
\]  
(295)

We now use Hölder’s inequality with indexes \( \frac{q}{p'}, \frac{q}{q-p} \) and control the moment term with Proposition 4 to find that, for some new \( \alpha > 0 \),
\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \mu_{t}^{N,K}, \phi_t^K(\mu) \right) \right] \leq C \mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} \mathcal{W}_Y \left( \mu_{t}^{N,K}, \phi_t^K(\mu_0^{N,K}) \right) \right]^{\alpha} \mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} \Lambda_{q} \left( \mu_{t}^{N,K}, \phi_t^K(\mu_0^{N,K}) \right) \right] 
\leq C a N^{-\alpha \eta} \exp(C a K (1 + t_{\text{fin}})) \cdot C a (1 + t_{\text{fin}}). 
\]  
(296)

Absorbing constants and the moment factors into the exponent, we have shown that, for some \( \alpha = \alpha(p, q, d) > 0 \),
\[
\mathbb{E} \left[ \sup_{t \leq t_{\text{fin}}} T_p \left( \mu_{t}^{N,K}, \phi_t^K(\mu) \right) \right] \leq N^{-\alpha} \exp(C a K (1 + t_{\text{fin}})).
\]  
(297)

The conclusion now follows by comparing \( \phi_t^K(\mu_0^{N,K}) \) and \( \phi_t^K(\mu_0) \) using Lemma 3. \( \square \)

### B. Appendix: Proof of Propositions 1, 2, 3

We finally address the wellposedness issues regarding the labelled and unlabelled Kac processes, which have been deferred. We will now prove Propositions 1, which describes the relationships between the labelled and unlabelled dynamics, and Propositions 2, 3, which assert a moderate wellposedness for the stochastic differential equation (LK) and of the martingale problem for the generator (5) of the unlabelled dynamics. Our strategy is as follows: the first item of Proposition 1 is elementary, and relies on a consistency between the unlabelled and labelled generators \( \mathcal{G}, \mathcal{G}^L \); for the second item, we carefully state a result of Kurtz [27, 28] and show how it applies in our case. For Proposition 2, we can show existence by standard techniques for martingale problems, using tightness and consistency of the generators; this does not use any result in the paper, and can be read independently of the more delicate estimates. For uniqueness in Proposition 3, we use the coupling and estimates in Section 5, which we emphasise do not rely on this result. We do not seek any estimates uniformly in \( N \), and we can replace moment estimates with the trivial bound \( |V_t| \leq \sqrt{N} \). For ease of presentation, we will use the estimates we have already developed in this paper, although those from the literature [20] would work equally well.
Let us recall some notation which will be needed. We will frequently move between objects defined on the labelled Kac sphere

\[ S_N = \left\{ \mathcal{V}^N = (V^1, \ldots, V^N) \in (\mathbb{R}^d)^N, \sum_{i=1}^N V^i = 0, \sum_{i=1}^N |V^i|^2 = N \right\} \]  

(298)

and the unlabelled state space \( \mathcal{S}_N \); we recall that \( \theta_N : \mathcal{S}_N \to \mathcal{S}_N \) is the map

\[ \mathcal{V}^N = (V^1, \ldots, V^N) \mapsto \frac{1}{N} \sum_{i=1}^N \delta_{V^i}. \]  

(299)

For clarity, we will indicate functions on \( \mathcal{S}_N \) with a \( \widehat{\cdot} \) to distinguish them from those on \( \mathcal{S}_N \). We will equip \( \mathcal{S}_N \) with the distance

\[ |\mathcal{V}^N - \mathcal{W}^N| := \sum_{i=1}^N |V^i - W^i| \]  

(300)

where the right-hand side is the Euclidean norm on \( \mathbb{R}^d \). We will write \( \mathcal{W}_1^1, \infty (\mathcal{S}_N) \) for the Sobolev space of functions \( \widehat{F} : \mathcal{S}_N \to \mathbb{R} \) which are Lipschitz with respect to this distance, equipped with the norm

\[ \| \widehat{F} \|_{W^1, \infty (\mathcal{S}_N)} := \max \left( \sup_{\mathcal{V}^N} |\widehat{F}(\mathcal{V}^N)|, \sup_{\mathcal{V}^N \neq \mathcal{W}^N} \frac{|\widehat{F}(\mathcal{V}^N) - \widehat{F}(\mathcal{W}^N)|}{|\mathcal{V}^N - \mathcal{W}^N|} \right) \]  

(301)

and define \( \mathcal{W}_1^1 (\mathcal{S}_N) \) similarly, equipping \( \mathcal{S}_N \) with the Wasserstein1 distance \( \mathcal{W}_1 \). It is elementary to show that these spaces are separable. Let us also recall, for convenience, the generators of the labelled and unlabelled dynamics, given respectively by

\[ (G^L F)(\mu_N) = N \int_{\mathbb{R}^d \times \mathbb{R}^d \times S^d} (F(\mu_N, v, v_*, \sigma) - F(\mu^N)) B(v - v_*, \sigma) \mu_N (dv) \mu_N (dv_*) d\sigma; \]  

(302)

\[ (G^L \widehat{F})(\mathcal{V}^N) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \int_{S^d-1} \left( \widehat{F}(\mathcal{V}^N_{i,j,\sigma}) - \widehat{F}(\mathcal{V}^N) \right) d\sigma \]  

(303)

for Lipschitz functions \( F \in \mathcal{W}_1^1 (\mathcal{S}_N) \), \( \widehat{F} \in \mathcal{W}_1^1 (\mathcal{S}_N) \) respectively. With this notation fixed, we turn to the proof of the two propositions.

**Proof of Proposition 1.** For item i), observe the following consistency between the unlabelled generator (302) and labelled generator (303), which follows from the Sym\((N)\) symmetry of the labelled dynamics: if \( F \in \mathcal{W}_1^1 (\mathcal{S}_N) \), then \( \widehat{F} := F \circ \theta_N \in \mathcal{W}_1^1 (\mathcal{S}_N) \), and

\[ G^L (F \circ \theta_N) = (G^L F) \circ \theta_N. \]  

(304)

Now, let \( \mathcal{V}^N_t \) be a labelled Kac process, for some filtration \( (\mathcal{F}_t)_{t \geq 0} \); it follows that \( \mathcal{V}^N_t \) solves the martingale problem for (303) for the same filtration. Now, let \( \mu^N_t = \theta_N(\mathcal{V}^N_t) \) be the
associated empirical measures, and fix $F \in W^{1,\infty}(S_N)$. For $\tilde{F} = F \circ \theta_N$ as above, the consistency (304) gives

$$F(\mu_i^N) - F(\mu_0^N) - \int_0^t (\mathcal{G}F)(\mu_s^N)ds = \tilde{F}(\nu_i^N) - \tilde{F}(\nu_0^N) - \int_0^t (\mathcal{G}^{L}\tilde{F})(\nu_s^N)ds. \quad (305)$$

The right-hand side is a martingale by assumption, and hence $\mu_i^N$ solves the martingale problem for (302) in the filtration $(\mathcal{F}_t)_{t \geq 0}$, as desired; in particular, $\mu_i^N$ is a Markov process with generator (302).

For item ii), we will use the next result, which generalises the implication needed, due to Kurtz [27,28]. Let us first fix some terminology. For a topological space $E$, let us write $\overline{C}(E)$ for the space of bounded, continuous functions on $E$, $\mathcal{B}(E)$ for the space of bounded, $\mathcal{B}$-measurable functions on $E$, and $\mathcal{P}(E)$ for the space of Borel probability measures. Given another such space $E_0$, a transition function $\alpha$ from $E_0$ to $E$ is a mapping from $E_0 \to \mathcal{P}(E)$ such that, for all Borel sets $A \subset E$, the map $y \mapsto \alpha(y, A)$ is a Borel function on $E_0$; for such $\alpha$ and $f \in \mathcal{B}(E)$, define $\alpha f \in \mathcal{B}(E_0)$ by

$$(\alpha f)(y) := \int_E f(z)\alpha(y, dz). \quad (306)$$

We will write $M_E[0, \infty), D_E[0, \infty)$ for the measurable, respectively càdlàg functions from $[0, \infty)$ to $E$.

Let us say that a linear operator $A \subset B(E) \times B(E)$ is separable if there exists a countable subset $\{f_\beta, \beta \geq 1\} \subset \mathcal{D}(A)$ such that, for all $(f, g) \in A$, there exists a subsequence $\beta_i \to \infty$ such that $(f_i, A f_i)$ are bounded uniformly in $i$, and converge pointwise to $(f, g)$.

We say that a linear operator $A$ is a pre-generator if it is dissipative, and there exists a sequence of functions $q_n : E \to \mathcal{P}(E), r_n : E \to [0, \infty)$ such that, for all $f \in \mathcal{D}(A)$, we have the pointwise convergence

$$r_n(x) \int_E (f(y) - f(x))q_n(x, dy) \to (Af)(x) \quad \text{for all } x \in E. \quad (307)$$

With these definitions, we can state the following result, which appears as part of [28, Theorem 1.4]:

**Proposition 10.** Let $(E, r), (E_0, r_0)$ be complete, separable metric spaces. Let $A \subset \overline{C}(E) \times \overline{C}(E)$ be a linear operator which is separable and a pre-generator, and whose domain $\mathcal{D}(A)$ separates points in $E$. Suppose that $\theta : E \to E_0$ is Borel measurable, and $\alpha$ is a transition function from $E_0$ to $E$ satisfying the compatibility condition $\alpha(y, \theta^{-1}(y)) = 1$ for all $y \in E_0$. Let $A^\theta$ be the linear operator

$$A^\theta = \{ (\alpha f, \alpha(A f)) : f \in \mathcal{D}(A) \} \subset B(E_0) \times B(E_0). \quad (308)$$

Let $\mathcal{L}_0 \subset \mathcal{P}(E_0)$, and let $\tilde{\mathcal{L}}_0 = \alpha_\# \mathcal{L}_0 \subset \mathcal{P}(E)$ be given by

$$\tilde{\mathcal{L}}_0(A) = \int_{E_0} \alpha(y, A)\mathcal{L}_0(dy). \quad (309)$$

If $\tilde{\mu} = (\tilde{\mu}_t)_{t \geq 0}$ is a solution of the martingale problem for $(A^\theta, \mathcal{L}_0)$, then there exists a solution $\mathcal{V}$ of the martingale problem for $(A, \tilde{\mathcal{L}}_0)$ such that $\tilde{\mu}$ has the same law on $M_{E_0}[0, \infty)$ as $\mu = \theta \circ \mathcal{V}$. Further, if $\tilde{\mu}$, and hence $\mu$, has a modification with sample paths in $D_{E_0}[0, \infty)$, then the modified $\tilde{\mu}, \mu$ have the same law on $D_{E_0}[0, \infty)$. 
Let us now show how this applies in our case. We will take \( E, E_0 \) to be the labelled and unlabelled Kac spheres \( E = \tilde{S}_N, E_0 = S_N \) respectively, equipped with the metrics as above. We take \( \mathcal{A} \) to be the labelled generator \( G^L \) given by (27), defined on \( F \in W^{1, \infty}(S_N) \), and let \( \theta = \theta_0 \) be given by (299). We define \( \alpha \) as the average over the preimage

\[
\alpha(\mu^N) = \frac{1}{\#\theta_0^{-1}(\mu^N)} \sum_{\nu^N \in \theta_0^{-1}(\mu^N)} \delta_{\nu^N}. \tag{310}
\]

We remark that, if \( \mu^N \in S_N \) and \( \nu^N \in \theta_0^{-1}(\mu^N) \), then \( \alpha(\mu^N) \) can be rewritten

\[
\alpha(\mu^N) = \frac{1}{N!} \sum_{\pi \in \text{Sym}(N)} \delta_{\nu^N, \pi} \tag{311}
\]

where \( \nu^N, \pi \) denotes the action of \( \pi \in \text{Sym}(N) \) permuting the \( N \) components \( V^1, ..., V^N \in \mathbb{R}^d \) of \( \nu^N \). It is elementary, if somewhat tedious, to check that with these choices, the linear operator \( \mathcal{A}^\theta \) is exactly the unlabelled generator \( G \), defined on \( W^{1, \infty}(S_N) \); the inclusion \( \mathcal{G} \subset \mathcal{A}^\theta \) is exactly the statement (304), and for the other inclusion \( \mathcal{A}^\theta \subset \mathcal{G} \), we use (311) to check that, for \( F : S_N \rightarrow \mathbb{R} \) Lipschitz, \( \alpha \tilde{F} : S_N \rightarrow \mathbb{R} \) is Lipschitz, and straightforward calculations show that \( G(\alpha \tilde{F}) = \alpha(G^L \tilde{F}) \) as desired.

To see that \( \mathcal{A} = G^L \) is separable, we note that \( W^{1, \infty}(S_N) \) is separable, and \( G^L : W^{1, \infty}(S_N) \rightarrow L^{\infty}(S_N) \) is a bounded linear map. Its graph is therefore separable in the stronger topology induced by \( G^L \subset W^{1, \infty}(S_N) \times L^{\infty}(S_N) \), and so is separable in the topology of bounded pointwise convergence in the definition above.

To see that \( G^L \) is a pregenerator, let us define \( G^L_K \) to be the cutoff equivalent, replacing \( B \) by the cutoff kernel \( B_K \) (6). It is straightforward to write \( G^L_K \) in the form desired, and \( G^L_K \rightarrow G_K \) in the space of bounded linear maps \( B(W^{1, \infty}(S_N), L^{\infty}(S_N)) \). Elementary, each \( G^L_K \) is the generator of a cutoff, labelled Kac process, and so generates a semigroup of contraction mappings; by the Lumer-Phillips Theorem, they are therefore dissipative; we can then take a limit to conclude that \( G^L \) is dissipative, and so is a a pregenerator.

We can now apply the conclusion of Proposition 10 above. Let us fix \( \mu_0^N \in S_N \), and let \((\tilde{\mu}_i^N)_{i \geq 0} \) be a solution to the martingale problem for the unlabelled generator (5) starting at \( \mu_0^N \). The law \( \tilde{\mathcal{L}}_0 \) given by Proposition 10 exactly corresponds to picking \( V_0^N \in \theta_0^{-1}(\mu_0^N) \) uniformly at random, as in the statement of the proposition, and by the result quoted above, there exists a solution to the martingale problem for (27), starting at \( V_0^N \) such that \( \tilde{\mu}_i^N \) has the same law as \( \theta_0(V_i^N) \). \( V_i^N \) is therefore a weak solution to the stochastic differential equation (LK), and so we have proven the claim of item ii).

**Proof of Proposition 2.** Let us fix \( V_0^N \); for each \( K, V_i^N, K \) be a solution to (eLK), starting at \( V_0^N \), with cutoff parameter \( K \). Since the rates are finite, such processes can be constructed elementarily, and have uniqueness in law. We check tightness via Aldous’ criterion; thanks to the energy constraint, each \( V_i^N, K \) takes values in \([-N^{1/2}, N^{1/2}]^d \), and for equicontinuity, we estimate

\[
\int_0^\infty dz \int_{[\mathbb{S}^{d-2}]} d\varphi |\alpha(v, v_\bullet, z, \varphi)| \leq C|v - v_\bullet|^{1+\gamma} \leq C(1+|v|^2 + |v_\bullet|^2). \tag{312}
\]

As above, let \( G^L, G^L_K \) be the (noncutoff/cutoff) labelled generators, and fix \( \tilde{F} \in W^{1, \infty}(S_N) \). As mentioned above, is straightforward to show that \( G^L_K \tilde{F} \) is continuous, and converge uniformly to \( G^L \tilde{F} \); it follows that any subsequential limit point of \( V_i^N, K \), as \( K \rightarrow \infty \), is a solution to the martingale problem for (27), and hence is a weak solution to (LK). \( \Box \)
Proof of Proposition 3. For uniqueness in law, let $V^N_t$ be any solution to (LK) starting at $V^N_0$. We now apply Lemma 8; fix $p > p_0(G, d), K > K_0(G, p, d)$ as in the statement, and $0 \leq t_1 < \ldots < t_m$, we take $b = N^{(p+\gamma)/2}$, so that $T^N_{b} = T^N_{b,K} = \infty$. The cited lemma now shows that $(V^N_{t_i})_{i \leq m}$ is the limit in probability, of $(V^N_{t_i})_{i \leq m}$, for cutoff labelled Kac processes $V^N_{t,K}$ starting at $V^N_0$, as $K \to \infty$. Since the law of each $V^N_{t,K}$ is uniquely determined, the same is true of the $m$-tuple $(V^N_{t_i})_{i \leq m}$. Since $t_i$ were arbitrary, we conclude that the law of $V^N_t$ is unique, as claimed. \qed

References

1. Alexandre, R.: Remarks on 3D Boltzmann linear equation without cutoff. Transp. Theory Stat. Phys. 28(5), 433–473, 1999
2. Alexandre, R., Desvillettes, L., Villani, C., Wennberg, B.: Entropy dissipation and long-range interactions. Arch. Ration. Mech. Anal. 152(4), 327–355, 2000
3. Arkeryd, L.: On the Boltzmann equation. In: Archive for Rational Mechanics and Analysis (1972)
4. Bolley, F., CARRILLO, J.A.: Tanaka theorem for inelastic Maxwell models. Commun. Math. Phys. 276(2), 287–314, 2007
5. Cortez, R., Fontbona, J.: Quantitative uniform propagation of chaos for Maxwell molecules. Commun. Math. Phys. 357(3), 913–941, 2018
6. Darling, R.W.R., Norris, J.R.: Differential equation approximations for Markov chains. Probab. Surv. 5, 37–79, 2008
7. Desvillettes, L.: Some applications of the method of moments for the homogeneous Boltzmann and Kac equations. Arch. Ration. Mech. Anal. 123(4), 387–404, 1993
8. Desvillettes, L., Graham, C., MÉLÉARD, S.: Probabilistic interpretation and numerical approximation of a Kac equation without cutoff. Stoch. Process. Appl. 84(1), 115–135, 1999
9. Desvillettes, L., Wennberg, B.: Smoothness of the solution of the spatially homogeneous Boltzmann equation without cutoff. Comm. Partial. Diff. Eq. 29(1–2), 133–155, 2005
10. Desvillettes, L., Mouhot, C.: Stability and uniqueness for the spatially homogeneous Boltzmann equation with long-range interactions. Arch. Ration. Mech. Anal. 193(2), 227–253, 2009
11. Elmroth, T.: Global boundedness of moments of solutions of the Boltzmann equation for forces of infinite range. Arch. Ration. Mech. Anal. 82(1), 1–12, 1983
12. Escobedo, M., Mischler, S.: Scalings for a ballistic aggregation equation. J. Stat. Phys. 141(3), 422–458, 2010
13. Fournier, N., MÉLÉARD, S.: A Markov process associated with a Boltzmann equation without cutoff and for non-Maxwell molecules. J. Stat. Phys. 104(1), 359–385, 2001
14. Fournier, N., MÉLÉARD, S.: A stochastic particle numerical method for 3D Boltzmann equations without cutoff. Math. Comput. 71(238), 583–604, 2002
15. Fournier, N.: Uniqueness for a class of spatially homogeneous Boltzmann equations without angular cutoff. J. Stat. Phys. 125(4), 923–942, 2006
16. Fournier, N., Guérin, H.: On the uniqueness for the spatially homogeneous Boltzmann equation with a strong angular singularity. J. Stat. Phys. 131(4), 749–781, 2008
17. Fournier, N., Mouhot, C.: On the well-posedness of the spatially homogeneous Boltzmann equation with a moderate angular singularity. Commun. Math. Phys. 289(3), 803–824, 2009
18. Fournier, N., Guillin, A., From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules. In: Annales Scientifiques de l’École Normale Supérieure, vol. 50, no. 1, pp. 157–199 (2017)
19. Fournier, N., Hauray, M.: Propagation of chaos for the Landau equation with moderately soft potentials. *Ann. Probab.* **44**(6), 3581–3660, 2016
20. Fournier, N., Mischler, S.: Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules. *Ann. Probab.* **44**(1), 589–627, 2016
21. Fournier, N.: On exponential moments of the homogeneous Boltzmann equation for hard potentials without cutoff. *Commun. Math. Phys.* **387**(2), 973–994, 2021
22. Fournier, N., Heydecker, D.: Stability, well-posedness and regularity of the homogeneous Landau equation for hard potentials. In: *Annales de l’Institut Henri Poincaré C, Analyse non linéaire*. Elsevier Masson (2021)
23. Graham, C., Méleard, S.: Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *Ann. Probab.* **25**, 115–132, 1997
24. Heydecker, D.: Pathwise convergence of the hard spheres Kac process. *Ann. Appl. Probab.* **29**(5), 3062–3127, 2019
25. Heydecker, D.: Kac’s process and some probabilistic aspects of the Boltzmann equation. Doctoral dissertation, University of Cambridge (2021).
26. Kac, M.: Foundations of kinetic theory. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*, vol. III, pp. 171–197 (1956)
27. Kurtz, T.: Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.* **3**, 1–29, 1998
28. Kurtz, T.G., Equivalence of stochastic equations and martingale problems. In: *Stochastic analysis 2010*, pp. 113–130. Springer Berlin, Heidelberg (2011)
29. Lu, X.: Conservation of energy, entropy identity, and local stability for the spatially homogeneous Boltzmann equation. *J. Stat. Phys.* **96**(3), 765–796, 1999
30. Lu, X., Mouhot, C.: On measure solutions of the Boltzmann equation, part I: moment production and stability estimates. *J. Differ. Equ.* **252**(4), 3305–3363, 2012
31. McKean, H.P.: Fluctuations in the kinetic theory of gases. *Commun. Pure Appl. Math.* **28**(4), 435–455, 1975
32. Mischler, S., Wennberg, B.: On the spatially homogeneous Boltzmann equation. In: *Annales de l’Institut Henri Poincaré C, Analyse non linéaire*, vol. 16, no. 4, pp. 467–501. Elsevier Masson (1999)
33. Mischler, S., Mouhot, C.: Kac’s program in kinetic theory. *Inventiones mathematicae* **193**(1), 1–147, 2013
34. Norris, J.: A consistency estimate for Kac’s model of elastic collisions in a dilute gas. *Ann. Appl. Probab.* **26**(2), 1029–1081, 2016
35. Povzner, A.Y.: On the Boltzmann equation in the kinetic theory of gases. *Matematiceskii Sbornik* **100**(1), 65–86, 1996
36. Pulvirenti, A., Toscani, G.: The theory of the nonlinear Boltzmann equation for Maxwell molecules in Fourier representation. *Annali di Matematica Pura ed Applicata* **171**(1), 181–204, 1996
37. Rousset, M.: A N-uniform quantitative Tanaka’s theorem for the conservative Kac’s N-particle system with Maxwell molecules. arXiv preprint arXiv:1407.1965 (2014)
38. Salem, S.: Propagation of chaos for the Boltzmann equation with moderately soft potentials. arXiv preprint arXiv:1910.01883 (2019)
39. Sznitman, A.S.: Topics in propagation of chaos. In: *Ecole d’Été de probabilités de Saint-Flour XIX–1989*, pp. 165–251. Springer, Berlin (1991)
40. Tanaka, H.: Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **46**(1), 67–105, 1978
41. Tanaka, H.: On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules. In: *Stochastic Processes: Selected Papers of Hiroshi Tanaka*, pp. 101–117 (2002)
42. Toscani, G., Villani, C.: Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas. *J. Stat. Phys.* **94**(3), 619–637, 1999
43. Wennberg, B.: Entropy dissipation and moment production for the Boltzmann equation. *J. Stat. Phys.* **86**(5), 1053–1066, 1997

 DANIEL HEYDECKER  
 Max Planck Institut für Mathematik in den Naturwissenschaften,  
 Inselstraße 22,  
 04103 Leipzig  
 Germany.  
e-mail: daniel.heydecker@mis.mpg.de

(Received August 10, 2020 / Accepted February 21, 2022)  
Published online April 12, 2022  
© The Author(s) (2022)