



# Quantitative Uniform Propagation of Chaos for Maxwell Molecules

Roberto Cortez<sup>1</sup>, Joaquin Fontbona<sup>2</sup> 

<sup>1</sup> CIMFAV, Facultad de Ingeniería, Universidad de Valparaíso, General Cruz 222, Valparaíso, Chile.  
E-mail: [rcortez@dim.uchile.cl](mailto:rcortez@dim.uchile.cl)

<sup>2</sup> Department of Mathematical Engineering and Center for Mathematical Modeling, UMI(2807) UCHILE-CNRS, University of Chile, Casilla 170-3, Correo 3, Santiago, Chile. E-mail: [fontbona@dim.uchile.cl](mailto:fontbona@dim.uchile.cl)

Received: 1 January 2016 / Accepted: 24 December 2017

Published online: 17 February 2018 – © Springer-Verlag GmbH Germany, part of Springer Nature 2018

**Abstract:** We prove propagation of chaos at explicit polynomial rates in Wasserstein distance  $\mathcal{W}_2$  for Kac's  $N$ -particle system associated with the spatially homogeneous Boltzmann equation for Maxwell molecules. Our approach is mainly based on novel probabilistic coupling techniques. Combining them with recent stabilization results for the particle system we obtain, under suitable moments assumptions on the initial distribution, a uniform-in-time estimate of order almost  $N^{-1/3}$  for  $\mathcal{W}_2^2$ .

## 1. Introduction and Main Results

*1.1. The Boltzmann equation.* The *spatially homogeneous Boltzmann equation* predicts that the density  $f_t(v)$  of particles with velocity  $v \in \mathbb{R}^3$  at time  $t \geq 0$  in a spatially homogeneous dilute gas subjected to binary collisions, satisfies

$$\partial_t f_t(v) = \frac{1}{2} \int_{\mathbb{R}^3} dv_* \int_{\mathbb{S}^2} d\sigma [f_t(v') f_t(v'_*) - f_t(v) f_t(v_*)] B(|v - v_*|, \theta), \quad (1)$$

where  $v'$  and  $v'_*$  are the *pre-collisional velocities*, given by

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma, \quad v'_* = \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma, \quad (2)$$

and  $\theta$  is the *deviation angle*, defined by  $\cos \theta = \sigma \cdot (v - v_*) / |v - v_*|$ . The *collision kernel*  $B(|v - v_*|, \theta) \geq 0$  describes the rate at which collisions between pairs of particles occur, and depends on the type of physical interactions among them. Solutions of (1) preserve mass  $\int_{\mathbb{R}^3} f_t(v) dv$ , momentum  $\int_{\mathbb{R}^3} v f_t(v) dv$  and kinetic energy  $\int_{\mathbb{R}^3} |v|^2 f_t(v) dv$ , so we may and shall assume that  $\int_{\mathbb{R}^3} f_t(v) dv = 1$  for all  $t \geq 0$ .

Equation (1) has been extensively studied for several decades. We refer the reader to Cercignani [4] for physical background on the Boltzmann equation and to Villani [31], Alexandre [1] and Mischler and Mouhot [22] for historical accounts on aspects of its mathematical theory.

Typically, one assumes that  $B : \mathbb{R}_+ \times (0, \pi] \rightarrow \mathbb{R}_+$  has the form

$$B(z, \theta) \sin \theta = z^\gamma \beta(\theta),$$

for some  $\gamma \in (-3, 1]$ , and some function  $\beta : (0, \pi] \rightarrow \mathbb{R}_+$  which, for symmetry reasons, can be taken to be equal to 0 on  $(\pi/2, \pi]$ . In this paper, we assume that  $\gamma = 0$  and  $\beta(\theta) \sim \theta^{-1-\nu}$  near 0 for  $\nu = 1/2$ , a setting referred to as *Maxwell molecules* case.

*1.2. Particle system and propagation of chaos.* As a step to rigorously justify the interpretation of the Boltzmann equation as a representation of the evolution of a very large number of interacting particles, Kac [19] suggested to study the limit, as  $N$  goes to  $\infty$ , of some exchangeable stochastic system of  $N$  of such particles, defined as a continuous-time pure-jump Markov process on  $(\mathbb{R}^3)^N$ . For a simplified one dimensional version of the nonlinear equation (1), he in fact proved that if the joint law of  $k$  particles at time 0 weakly converges as  $N \rightarrow \infty$  to the  $k$ -fold product of an initial density  $f_0$  in  $\mathbb{R}$ , then the same holds true at times  $t > 0$ , with  $f_t$  as the limit density.

This property, termed *propagation of chaos*, is equivalent to the convergence of the empirical measure of the system at each time  $t$  to the solution of the nonlinear equation, and has been established, under different convergence criteria, for a wide class of models including the true Boltzmann equation (1). For general background on propagation of chaos, we refer the reader to Sznitman [27], Méléard [21] and Mischler and Mouhot [22] (see also Sect. 1.4 below for historical and recent references).

More specifically, we consider the particle system given by the  $(\mathbb{R}^3)^N$ -valued Markov process with infinitesimal generator  $\mathcal{A}^N$  defined as follows: for all Lipschitz bounded function  $\Phi : (\mathbb{R}^3)^N \rightarrow \mathbb{R}$  and  $\mathbf{v} = (v^1, \dots, v^N) \in (\mathbb{R}^3)^N$ ,

$$\mathcal{A}^N \Phi(\mathbf{v}) = \frac{1}{2(N-1)} \sum_{i \neq j} \int_{\mathbb{S}^2} d\sigma [\Phi(\mathbf{a}_{ij}(\mathbf{v}, \sigma)) - \Phi(\mathbf{v})] B(\theta), \tag{3}$$

where  $B(\theta) \sin \theta := \beta(\theta)$ , and  $\mathbf{a}_{ij}(\mathbf{v}, \sigma) \in (\mathbb{R}^3)^N$  is the vector  $\mathbf{v}$  with its  $i$ -th and  $j$ -th  $\mathbb{R}^3$ -valued components respectively replaced by  $v'$  and  $v'_*$  given by (2) with  $v = v^i$  and  $v_* = v^j$ . Particles start with a symmetric law  $G_0^N$  on  $(\mathbb{R}^3)^N$ . We denote<sup>1</sup> this stochastic interacting particle system by  $\mathbf{V}_t = (V_t^1, \dots, V_t^N)$ .

Hence, any pair of particles  $i$  and  $j$  with velocities  $v = V_t^i$  and  $v_* = V_t^j$  interact with deviation angle  $\theta$  at rate  $\beta(\theta)/2(N-1)$ , and then they update their velocities to  $v'$  and  $v'_*$  given by (2), with  $\sigma \in \mathbb{S}^2$  uniformly chosen at random among unitary vectors such that  $\sigma \cdot (v - v_*)/|v - v_*| = \cos \theta$ ; notice that  $v'$  and  $v'_*$  as defined in (2) now appear in the role of *post-collisional velocities*, consistently with the weak form (9) of Eq. (1). The fact that the function  $\beta(\theta)$  has a non-integrable singularity at  $\theta = 0$  entails that particles jump infinitely many times on every finite time interval (all but finitely many jumps corresponding to very small deviation angles). One can check that the quantities  $\sum_{i=1}^N v^i$  and  $\sum_{i=1}^N |v^i|^2$ , corresponding to momentum and kinetic energy, are a.s. exactly preserved by the corresponding random dynamics.

<sup>1</sup> For notational simplicity, we do not make explicit the dependence of the system on  $N$ .

An explicit construction of the system  $\mathbf{V}$  will be given in Lemma 8, Sect. 2. Let us mention for the moment that, under the assumptions we will make, a unique (in law) Markov process with càdlàg trajectories and generator (3) will exist, for each value of  $N \in \mathbb{N}$ .

The goal of the present paper is to establish a fully explicit, uniform in time propagation of chaos rate, for the Kac  $N$ -particle system  $\mathbf{V}$ . We adopt here a probabilistic pathwise approach, as pioneered by Tanaka [28, 29] and Sznitman [27]. The main idea is to extend the coupling techniques for binary-jump particle systems introduced in [5] to the much more difficult framework of the Boltzmann equation. We will also rely on the analytic approach and estimates of Fournier and Mischler [13]. Moreover, combining these ideas with a uniform-in- $N$  equilibration result, recently established by Rousset [25] for Kac particles in the Maxwell molecules case, we will obtain the sharpest propagation of chaos estimates in Wasserstein distance so far available in this setting.

*1.3. Main results.* Let us first fix some notation and specify our hypotheses. Given a metric space  $E$ ,  $p \geq 1$  and  $k \in \mathbb{N}$ , let  $\mathcal{P}(E)$ ,  $\mathcal{P}_p(E)$  and  $\mathcal{P}_p^{\text{sym}}(E^k)$  respectively denote the space of probability measures on  $E$ , the subspace of probability measures on  $E$  with finite  $p$ -moment, and the subspace of  $\mathcal{P}_p(E^k)$  consisting of symmetric probability measures on  $E^k$  with finite  $p$ -moment. Given a vector  $\mathbf{x} = (x^1, \dots, x^N) \in (\mathbb{R}^3)^N$ , we define the empirical measures  $\bar{\mathbf{x}} \in \mathcal{P}(\mathbb{R}^3)$  and  $\bar{\mathbf{x}}^i \in \mathcal{P}(\mathbb{R}^3)$  for any  $i \in \{1, \dots, N\}$ , by

$$\bar{\mathbf{x}} := \frac{1}{N} \sum_{j=1}^N \delta_{x^j} \quad \text{and} \quad \bar{\mathbf{x}}^i := \frac{1}{N-1} \sum_{j \neq i} \delta_{x^j}. \tag{4}$$

The empirical measure of the particle system at time  $t \geq 0$  is thus denoted by  $\bar{\mathbf{V}}_t$ . Also, given a (mainly exchangeable) random vector  $\mathbf{X}$  on  $(\mathbb{R}^3)^N$ , we will denote its law by  $\mathcal{L}(\mathbf{X}) \in \mathcal{P}((\mathbb{R}^3)^N)$ , and the law of its  $k$  first components by  $\mathcal{L}^k(\mathbf{X}) \in \mathcal{P}((\mathbb{R}^3)^k)$ , for any  $k \leq N$ .

For  $\mu, \nu \in \mathcal{P}_2((\mathbb{R}^3)^k)$ , their *quadratic Wasserstein distance* is defined as

$$\mathcal{W}_2(\mu, \nu) = \inf_{\pi} \left( \int_{(\mathbb{R}^3)^k \times (\mathbb{R}^3)^k} |\mathbf{x} - \mathbf{y}|_k^2 \pi(d\mathbf{x}, d\mathbf{y}) \right)^{1/2} = \inf_{\mathbf{X}, \mathbf{Y}} \left( \mathbb{E} |\mathbf{X} - \mathbf{Y}|_k^2 \right)^{1/2},$$

where the first infimum is taken over all  $\pi \in \mathcal{P}_2((\mathbb{R}^3)^k \times (\mathbb{R}^3)^k)$  having marginals  $\mu$  and  $\nu$ , and the second infimum is taken over all random vectors  $\mathbf{X}$  and  $\mathbf{Y}$  on  $(\mathbb{R}^3)^k$  such that  $\mathcal{L}(\mathbf{X}) = \mu$  and  $\mathcal{L}(\mathbf{Y}) = \nu$ . Here we use the *normalized distance*  $|\cdot|_k$  on  $(\mathbb{R}^3)^k$  defined by  $|\mathbf{x}|_k^2 = \frac{1}{k} \sum_{i=1}^k |x^i|^2$ . It is known that the infimum is always reached, and a  $\pi$  attaining the first infimum or a pair  $(\mathbf{X}, \mathbf{Y})$  attaining the second one, is referred to as an *optimal coupling*.

The angular cross section function  $\beta$  will be assumed to satisfy

$$\exists \nu \in (0, 1), \quad c_0 \theta^{-1-\nu} \leq \beta(\theta) \leq c_1 \theta^{-1-\nu} \quad \forall \theta \in (0, \pi/2), \tag{5}$$

for some constants  $0 < c_0 < c_1$ . The initial distribution  $f_0$  will be assumed to satisfy

$$\exists p_0 > 4, \quad \int_{\mathbb{R}^3} |v|^{p_0} f_0(dv) < \infty. \tag{6}$$

Note that in the usual Maxwellian case, the quantities  $\nu$  and  $\gamma = 0$  are linked by the relation  $\nu = \frac{1-\gamma}{2}$ , but in our (slightly more general) context,  $\nu \in (0, 1)$  will be viewed as an independent parameter.

We now state our main results. See Definition 3 below for the notion of weak solutions of (1) that we will use, and Theorem 4 for its well-posedness .

**Theorem 1.** *Assume (5) and (6), and let  $(f_t)_{t \geq 0}$  be the unique weak solution of (1). Let  $G_0^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$  be given, and let  $\mathbf{V}_t = (V_t^1, \dots, V_t^N)$  be the particle system with generator (3) starting with law  $G_0^N$ . Then, there exists a constant  $C > 0$  such that for all  $t \geq 0$ ,*

$$\mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C \mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + C(1+t)^2 N^{-1/3}.$$

Note that one can simply choose  $G_0^N = f_0^{\otimes N}$  and then  $\mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) = 0$ , or assume that the term  $\mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$  goes to 0 at least as fast as  $N^{-1/3}$ . In either case, the previous theorem gives a chaos result in squared 2-Wasserstein distance for the Kac particle system associated to the Boltzmann equation for Maxwell molecules, with an explicit rate of order  $N^{-1/3}$ . The time dependence is quadratic.

It is worth noting that the proof of Theorem 1 also provides the same convergence rates for each  $k$ -marginal of the  $N$ -particle system defined by (3), when  $N$  goes to  $\infty$ , and that statements for cutoff systems can also be established (with similar dependence on the cutoff parameter as in [13]).

Under stronger conditions on the initial law, we obtain our most important result:

**Theorem 2.** *Under the same hypotheses of Theorem 1, assume additionally that  $\int \nu f_0(d\nu) = 0$ ,  $\int |\nu|^2 f_0(d\nu) = 1$ , and that  $R_p := \sup_N \mathbb{E}|V_0^1|^p < \infty$  for some  $p \geq 4$ . Then, there exists  $C > 0$  and for all  $\epsilon > 0$  there exists  $C_{p,\epsilon} > 0$  such that for all  $t \geq 0$ ,*

$$\mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C \mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + C_{p,\epsilon} N^{-(p-2)/3p+\epsilon}.$$

Thus, for each such  $p \geq 4$ , we obtain a uniform propagation of chaos estimate for Maxwell molecules in squared 2-Wasserstein distance at an explicit rate of order almost  $N^{-(p-2)/3p}$ , provided that  $\mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$  converges to 0 with the same rate or faster (one can simply take  $G_0^N = f_0^{\otimes N}$  as long as  $f_0$  has finite  $p$ -moment). For instance: if, together with (6), one only assumes  $R_p < \infty$  for  $p = 4$ , we obtain a chaos rate of order almost  $N^{-1/6}$ ; but if one assumes  $R_p < \infty$  for all  $p$ , then the rate is of order almost  $N^{-1/3}$ .

Let us mention that our techniques can be applied to the case in which  $\nu \in [1, 2)$  (implying the usual integrability condition  $\int_0^{\pi/2} \theta^2 \beta(\theta) < \infty$ ), but this would require additional technicalities in order to treat the probabilistic objects involved in the statements and proofs. To keep the exposition simple, we restrict ourselves to the case  $\nu \in (0, 1)$ , which includes the classical Maxwellian molecules ( $\nu = 1/2$ ). Let us further remark that our coupling techniques can be applied to Kac particles systems in the hard potentials and hard spheres cases as well, providing in those frameworks the same rate  $N^{-1/3}$  as in Theorem 1, but a much worse dependence on time. This will be addressed in forthcoming works.

*1.4. Comparison to known results and approaches.* The study of propagation of chaos for the Boltzmann equation was initiated in the paper [19] by Kac. Propagation of chaos

results were proved by several authors the decades thereafter, for different instances of the equation, in the weak convergence sense and without convergence rates. For instance, McKean [20] and Grünbaum [17] obtained such results for some models with bounded collision kernel, whereas the work of Sznitman [26] dealt with unbounded kernels in the hard spheres case. Tanaka [28,29] introduced a crucial tool in the probabilistic study of the propagation of chaos property, the *nonlinear process*, which represents the trajectory of a “typical particle” in the infinite population and whose time-marginal laws solve the Boltzmann equation. By coupling the particle system with independent nonlinear processes, Graham and Méléard [16] obtained one of the first quantitative propagation of chaos results which cover cutoff variants of the model. However, their approach, relying on total variation distance on path space, can not be extended to non-cutoff contexts, and provides bounds which increase exponentially in time.

More recently, in the remarkable work of Mischler and Mouhot [22], uniform-in-time propagation of chaos results in  $\mathcal{W}_1$  distance for Maxwellian molecules and hard spheres are established, with a slow (and hard to track) rate in  $N$ . Their method, of analytic nature, focused on the stability of the evolution of the time-marginal laws of the particle system and relied on the comparison between Wasserstein and other probability distances. Moreover, combining Theorem 5.1-(iii) of [22] with the estimates found in Step 3 of the proof of Theorem 8-(ii) of [3] by Carrapatoso, one can obtain a uniform-in-time chaos rate for Maxwellian molecules of order almost  $N^{-1/8}$  in some comparable distance, though only in the case of i.i.d. initial data conditioned to the Boltzmann sphere, and under a finite Fisher information assumption (this seems to be the best uniform rate so far available in the literature; we thank an anonymous referee for pointing out this to us).

In [13], using a coupling with independent nonlinear processes and optimal transport based techniques, Fournier and Mischler obtained a propagation of chaos result with an optimal rate of order  $N^{-1/2}$  in squared Wasserstein distance  $\mathcal{W}_2$  for *Nanbu* particle systems in the hard potentials, hard spheres and Maxwell molecules cases, in the latter setting with the same dependence on  $t$  as ours in Theorem 1. Analogous coupling arguments relying on optimal transport were developed earlier in Fontbona et al. [7] for Nanbu type diffusive approximations of the Landau equation, with less sharper explicit rates (due to the general coefficients and the suboptimal estimates for empirical measures of i.i.d. samples available by that time). Recall that, contrary to the Markov dynamics (3) of binary (also called *Bird type*) interactions, each particle in a Nanbu type system is driven by an independent noise source, which implies in the jump case that only one particle jumps at each collision. Such particle systems preserve momentum and energy only in mean, and hence are less meaningful from the physical point view (but are still relevant for numerical simulation purposes). Since the coupling constructions in [7,13] strongly relied on the independence of the noise sources for different particles, they cannot be applied to systems with true binary interactions like (3), where the random noises are shared.

In [5], we addressed this problem in the case of Kac’s one dimensional model and some generalizations. More precisely, we introduced a new coupling between an interacting particle system with effective binary interactions and a system of nonlinear processes driven by the same randomness sources, which thus turned out to be *not independent*. As part of that coupling argument, we had to show, in a second step, that these nonlinear processes become on their turn independent as  $N$  goes to  $\infty$ . With a similar strategy, the case of the Landau equation was recently addressed by Fournier and Guillin [10] in the hard potentials and Maxwell molecules cases. Relying also on a stabilization

result for the corresponding particle system (analogous to the one in [25]), they obtain in the latter case a uniform propagation of chaos estimate, similar to ours in the Boltzmann case.

One of the main additional difficulties that pathwise probabilistic approaches to the Boltzmann equation need to deal with, when compared to the one dimensional Kac model (and also to the Landau setting), is the lack of continuity of the parametrization of the collision angles, as a function of pre-collisional velocities. Here, we cope with this problem using optimal transport techniques and dealing with cutoff versions of the nonlinear process.

The assumptions on the initial distributions required in Theorem 2 are similar to those in [10], and are much more general than in all the available uniform propagation of chaos results for Maxwell molecules (no support constraint or regularity being needed). We notice that the rate of  $N^{-1/3}$  in  $\mathcal{W}_2^2$  for Bird type particle systems obtained here, in [5] and in [10], is slower than the  $N^{-1/2}$  rate valid for Nanbu type systems (corresponding to the optimal convergence rate in expected  $\mathcal{W}_2^2$  distance for the empirical measure of i.i.d. samples, established in [11]). An interesting question, raised in [5], is to what extent this sub-optimality is intrinsic to the interaction type, or a consequence of the techniques employed.

*1.5. Weak solutions and nonlinear processes.* Next we recall the notion of weak solutions for (1) we will work with, for which some definitions are needed. We follow [13]. Consider the function  $G : \mathbb{R}_+ \rightarrow (0, \pi/2]$  defined as  $G(z) = H^{-1}(z)$  where  $H : (0, \pi/2] \rightarrow \mathbb{R}_+$  is given by

$$H(\theta) := \int_{\theta}^{\pi/2} \beta(x) dx.$$

Consider also measurable functions  $\hat{i}, \hat{j} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  such that for every  $x \neq 0$ ,

$$\left( \frac{x}{|x|}, \frac{\hat{i}(x)}{|x|}, \frac{\hat{j}(x)}{|x|} \right)$$

is an orthonormal basis of  $\mathbb{R}^3$ . We may and shall assume that they are homogeneous functions, that is, one has  $\hat{i}(\lambda x) = \lambda \hat{i}(x)$  and  $\hat{j}(\lambda x) = \lambda \hat{j}(x)$  for all  $x \in \mathbb{R}^3$  and all  $\lambda \in \mathbb{R}$ . For  $v, v_* \in \mathbb{R}^3, \theta \in (0, \pi/2), \phi \in [0, 2\pi)$ , and  $z \geq 0$ , define

$$\begin{aligned} \Gamma(x, \phi) &:= (\cos \phi) \hat{i}(x) + (\sin \phi) \hat{j}(x), \\ a(v, v_*, \theta, \phi) &:= -\frac{1 - \cos \theta}{2} (v - v_*) + \frac{\sin \theta}{2} \Gamma(v - v_*, \phi), \\ c(v, v_*, z, \phi) &:= a(v, v_*, G(z), \phi). \end{aligned} \tag{7}$$

Note that when  $\phi$  varies in  $[0, 2\pi)$ , the vector  $v + a(v, v_*, \theta, \phi)$  ranges all over the circle centered at  $b = \frac{v+v_*}{2} + \cos \theta \frac{v-v_*}{2}$ , with radius  $r = \sin \theta \frac{|v-v_*|}{2}$  and orthogonal to  $d = \frac{v-v_*}{|v-v_*|}$ . Denote this circle by  $\mathcal{C}(v, v_*, \theta)$ , or alternatively by  $\tilde{\mathcal{C}}(b, r, d)$ .

These objects provide a suitable parametrization of the post-collisional velocities: it is straightforward to verify that for all  $v, v_* \in \mathbb{R}^3$  and for any Lipschitz bounded measurable function  $\Phi$  on  $\mathbb{R}^3$ ,

$$\begin{aligned}
 & \int_{\mathbb{S}^2} d\sigma [\Phi(v') - \Phi(v)] B(\theta) \\
 &= \int_0^{\pi/2} d\theta \int_0^{2\pi} d\phi [\Phi(v + a(v, v_*, \theta, \phi)) - \Phi(v)] \beta(\theta) \\
 &= \int_0^\infty dz \int_0^{2\pi} d\phi [\Phi(v + c(v, v_*, z, \phi)) - \Phi(v)],
 \end{aligned} \tag{8}$$

the integral being well defined since  $|a(v, v_*, \theta, \phi)| \leq C\theta|v - v_*|$  and  $\int_0^{\pi/2} \theta\beta(\theta)d\theta < \infty$  when (5) holds. By a slight abuse of notation, we still call  $v' = v + a(v, v_*, \theta, \phi)$ ,  $v'_* = v_* - a(v, v_*, \theta, \phi)$  and also  $v' = v + c(v, v_*, z, \phi)$ ,  $v'_* = v_* - c(v, v_*, z, \phi)$ .

**Definition 3.** We say that a collection  $(f_t)_{t \geq 0} \in C([0, \infty), \mathcal{P}_2(\mathbb{R}^3))$  is a *weak solution* for (1) if it preserves momentum and energy (that is,  $\int v f_t(dv) = \int v f_0(dv)$  and  $\int |v|^2 f_t(dv) = \int |v|^2 f_0(dv)$  for all  $t \geq 0$ ), and if for all bounded Lipschitz function  $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$  and for all  $t \geq 0$ ,

$$\begin{aligned}
 \int_{\mathbb{R}^3} \Phi(v) f_t(dv) &= \int_{\mathbb{R}^3} \Phi(v) f_0(dv) + \int_0^t ds \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} f_s(dv) f_s(dv_*) \int_0^{\pi/2} d\theta \beta(\theta) \\
 &\quad \times \int_0^{2\pi} d\phi [\Phi(v + a(v, v_*, \theta, \phi)) - \Phi(v)].
 \end{aligned} \tag{9}$$

The next statement provides the main analytical properties of Eq. (1) that we shall need. The proof of well-posedness can be found for instance in [30,31], whereas the proof of the existence of a density can be found in [9].

**Theorem 4.** Assume (5) and (6). Then, there exists a unique weak solution  $(f_t)_{t \geq 0} \in C([0, \infty), \mathcal{P}_2(\mathbb{R}^3))$  of (1), which satisfies  $\sup_t \int_{\mathbb{R}^3} |v|^{p_0} f_t(dv) < \infty$ . Moreover, if  $f_0$  is not a Dirac mass, then  $f_t$  has a density as soon as  $t > 0$ .

The nonlinear process, introduced by Tanaka in [28,29] to provide a probabilistic interpretation of the Boltzmann equation, can be defined in the present case through a stochastic integral equation with respect to some Poisson point measure. More specifically, consider the equation

$$dW_t = \int_0^\infty \int_0^{2\pi} \int_{\mathbb{R}^3} c(W_{t-}, v, z, \phi) \mathcal{M}(dt, dz, d\phi, dv), \tag{10}$$

where  $\mathcal{M}(dt, dz, d\phi, dv)$  is a Poisson point measure on  $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times \mathbb{R}^3$  with intensity  $dt dz d\phi f_t(dv)/2\pi$ . Following Tanaka’s ideas, under (5) and (6) Fournier and Méléard [12] proved weak existence and uniqueness in law for Eq. (10), together with the fact that  $(\mathcal{L}(W_t))_{t \geq 0}$  solves (9), hence  $\mathcal{L}(W_t) = f_t$  for all  $t$ . Any process having the same law as  $W$  is called a nonlinear process.

Unfortunately, we cannot carry out our coupling construction by relying only on weak existence of solutions to Eq. (10). This is why we will need to work with the *cutoff* nonlinear process instead. Given a cutoff level  $L > 0$ , this process  $W^L$  can be defined as the solution of a nonlinear SDE similar to (10), namely

$$dW_t^L = \int_0^\infty \int_0^{2\pi} \int_{\mathbb{R}^3} c_L(W_{t-}^L, v, z, \phi) \mathcal{M}^L(dt, dz, d\phi, dv), \tag{11}$$



where  $c_L(v, v_*, z, \phi) := c(v, v_*, z, \phi)\mathbf{1}_{\{z \leq L\}}$ . This time  $\mathcal{M}^L(dt, dz, d\phi, dv)$  is a Poisson point measure on  $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times \mathbb{R}^3$  with intensity  $dt dz d\phi f_t^L(dv)/2\pi$ , where  $(f_t^L)_{t \geq 0} \in C([0, \infty), \mathcal{P}_2(\mathbb{R}^3))$  is the unique solution to the nonlinear equation (9) with  $\beta$  replaced by  $\beta_L(z) := \beta(z)\mathbf{1}_{\{\theta \geq G(L)\}}$ , for which the well-posedness part of Theorem 4 applies. Strong well-posedness for (11) is straightforward: thanks to the indicator  $\mathbf{1}_{\{z \leq L\}}$ , the equation is nothing but a recursion for the values of  $W_t^L$  at its timely ordered jump-times. By standard arguments, it can be seen that any (weak or strong) solution to (11) satisfies  $\mathcal{L}(W_t^L) = f_t^L$  (more specifically, the collection  $(\mathcal{L}(W_t^L))_{t \geq 0}$  satisfies a linearized and cutoff version of (9), which in turn has  $(f_t^L)_{t \geq 0}$  as the unique solution, see for instance Theorem 3.1 of [12] for details). As expected, one can show that  $f_t^L \rightarrow f_t$  as  $L \rightarrow \infty$  (see Lemma 15 below).

1.6. *Idea of the proofs and plan of the paper.* To prove our results, following ideas introduced in [5], for each  $N$  and cutoff parameter  $L$  we will first couple in some optimal way, a suitable realization of the particle system  $\mathbf{V}_t$  with generator (3) (given below in (13)) with some system  $\mathbf{U}_t^L = (U_t^{1,L}, \dots, U_t^{N,L})$  of copies of the cutoff nonlinear process  $W^L$ . To do this, we will make use of optimal transport theory, in order to carefully construct the jumps of the system  $\mathbf{U}^L$ , in such a way that they mimic as closely as possible the jumps of the particle system  $\mathbf{V}$ . Roughly speaking, from this construction and Gronwall’s lemma, we will obtain an estimate like

$$\mathbb{E} \frac{1}{N} \sum_{i=1}^N |V_t^i - U_t^{i,L}|^2 \leq C \left[ \mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + (1+t)^2 \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{U}}_t^L, f_t^L) + tL^{1-2/\nu} \right]$$

for some constant  $C > 0$ . As mentioned earlier, the fact that we deal with a particle system with effective binary collisions will imply that the cutoff nonlinear processes  $U^{1,L}, \dots, U^{N,L}$  thus constructed are *not independent*. Therefore, in a similar way as in [5], we will need to “decouple”, in a second step, these processes, obtaining

$$\mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{U}}_t^L, f_t^L) \leq CN^{-1/3}$$

uniformly on  $L$  (see Lemma 12 and Corollary 14). We will then make  $L \rightarrow \infty$  to deduce the estimates of Theorem 1. Finally, from this and the results of [25], we will obtain the uniform-in-time chaos rate stated in Theorem 2.

In Sect. 2 we give an explicit construction of the particle system  $\mathbf{V}_t$  and, more importantly, we construct the coupling with the corresponding system  $\mathbf{U}_t^L$  of non-independent cutoff nonlinear processes used throughout this paper. In Sect. 3 we state and prove several technical results. In Sect. 4 we prove Theorem 1. Finally, in Sect. 5 we prove Theorem 2, along with some intermediate results that have interest on their own, such as the propagation of moments for the particle system without assuming bounded initial energy (see Corollary 17), and an equilibration result that extends the one by Rousset [25] (see Lemma 18).

## 2. Construction

In this section we explicitly construct the coupled system  $(\mathbf{V}, \mathbf{U}^L)$  in order to prove our results. These processes will be defined as solutions of stochastic integral equations driven by the same Poisson point measure. We follow [5, 13].



2.1. *The particle system.* Fix the number of particles  $N \in \mathbb{N}$ . We introduce the function  $\mathbf{i} : [0, N) \rightarrow \{1, \dots, N\}$  by  $\mathbf{i}(\xi) = \lfloor \xi \rfloor + 1$ , so that  $\mathbf{i}(\xi)$  is a discrete index associated to the continuous variable  $\xi$ . Let  $\mathcal{G} \subseteq \mathbb{R}^2$  be the set

$$\mathcal{G} = \{(\xi, \zeta) \in [0, N)^2 : \mathbf{i}(\xi) \neq \mathbf{i}(\zeta)\}.$$

Note that its area is  $|\mathcal{G}| = N(N-1)$ . Consider now a Poisson point measure  $\mathcal{N}(dt, dz, d\phi, d\xi, d\zeta)$  on  $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times [0, N) \times [0, N)$  with intensity

$$\frac{N}{2} dt dz \frac{d\phi}{2\pi} \frac{d\xi d\zeta \mathbf{1}_{\mathcal{G}}(\xi, \zeta)}{|\mathcal{G}|} = \frac{dt dz d\phi d\xi d\zeta \mathbf{1}_{\mathcal{G}}(\xi, \zeta)}{4(N-1)\pi}.$$

In words, the measure  $\mathcal{N}$  picks atoms  $(t, z) \in [0, \infty)^2$  with intensity  $\frac{N}{2} dt dz$  and for each such atom it also independently samples an angle  $\phi$  uniformly from  $[0, 2\pi)$  and a pair  $(\xi, \zeta)$  uniformly from the set  $\mathcal{G}$ . We will use the variables  $\xi$  and  $\zeta$  to choose indexes  $i = \mathbf{i}(\xi)$  and  $j = \mathbf{i}(\zeta)$  of the particles that interact at each jump. Additionally, given  $G_0^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$  and  $f_0 \in \mathcal{P}_2(\mathbb{R}^3)$  as in the statement of Theorem 1, we will in the sequel denote by

$$(\mathbf{V}_0, \mathbf{U}_0) \tag{12}$$

a realization, independent of  $\mathcal{N}$ , of the optimal coupling between  $G_0^N$  and  $f_0^{\otimes N}$ . Call  $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$  the complete, right-continuous filtration generated by  $(\mathbf{V}_0, \mathbf{U}_0)$  and  $\mathcal{N}$ . We denote by  $\mathbb{P}$  and  $\mathbb{E}$  the probability measure and expectation in the corresponding probability space.

We can now introduce the particle system  $\mathbf{V} = (V^1, \dots, V^N)$  as the solution, starting from the initial condition  $\mathbf{V}_0$ , of the stochastic equation

$$d\mathbf{V}_t = \int_0^\infty \int_0^{2\pi} \int_{[0, N)^2} \sum_{i \neq j} \mathbf{1}_{\{\mathbf{i}(\xi)=i, \mathbf{i}(\zeta)=j\}} \mathbf{c}_{ij}(\mathbf{V}_{t-}, z, \phi) \mathcal{N}(dt, dz, d\phi, d\xi, d\zeta) \tag{13}$$

where  $\mathbf{c}_{ij}(\mathbf{x}, z, \phi) \in (\mathbb{R}^3)^N$  is the vector with coordinates given by

$$(\mathbf{c}_{ij}(\mathbf{x}, z, \phi))^l = \begin{cases} c(x^i, x^j, z, \phi) & \text{if } l = i, \\ -c(x^j, x^i, z, \phi) & \text{if } l = j, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Weak existence and uniqueness of solutions for (13) holds, see Lemma 8 below.

Since we required  $\hat{i}$  and  $\hat{j}$  to be homogeneous functions, in particular they are odd functions and it can be easily seen that  $c(v, v_*, z, \phi) = -c(v_*, v, z, \phi)$ . Given a solution to (13), it follows that for each  $i = 1, \dots, N$ , the particle  $V^i$  satisfies the stochastic equation

$$dV_t^i = \int_0^\infty \int_0^{2\pi} \int_0^N c(V_{t-}^i, V_{t-}^{\mathbf{i}(\xi)}, z, \phi) \mathcal{N}^i(dt, dz, d\phi, d\xi), \tag{15}$$

where  $\mathcal{N}^i$  is given by

$$\mathcal{N}^i(dt, dz, d\phi, d\xi) := \mathcal{N}(dt, dz, d\phi, d\xi, [i-1, i)) + \mathcal{N}(dt, dz, d\phi, [i-1, i), d\xi).$$

That is,  $\mathcal{N}^i$  selects only the atoms of  $\mathcal{N}$  such that either  $\mathbf{i}(\xi) = i$  or  $\mathbf{i}(\zeta) = i$ . Clearly,  $\mathcal{N}^i$  is a Poisson point measure on  $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times [0, N)$  with intensity

$$\frac{dt dz d\phi d\xi \mathbf{1}_{A^i}(\xi)}{2(N-1)\pi},$$

where  $A^i := [0, N) \setminus [i - 1, i)$ . Thus, the term  $V_{t^-}^{i(\xi)}$  appearing in (15) is a  $\xi$ -realization of the (random) probability measure  $\tilde{V}_t^i$  (defined as in (4)). Therefore, from the point of view of the particle  $V^i$ , the dynamics is as follows:  $(t, z)$ -atoms are sampled with intensity 1 and for each such atom an angle  $\phi$  is chosen and a particle  $v_* = V_{t^-}^{i(\xi)}$  is selected at random among all the others;  $v = V_{t^-}^i$  (and  $v_* = V_{t^-}^{i(\xi)}$ ) then updates its state to  $v' = V_t^i$  (and  $v'_* = V_t^{i(\xi)}$ ) as given in (2).

2.2. *Coupling with a system of cutoff nonlinear processes.* The key observation is the following: in (15) with  $c_L$  in place of  $c$ , if one replaces  $V_{t^-}^{i(\xi)}$  by some realization of the probability measure  $f_t^L$ , then the resulting equation defines a cutoff nonlinear process as in (11). Moreover, we want to choose this  $f_t^L$ -distributed random variable in an optimal way (in the suitable sense), so that the resulting process remains close to  $V^i$ . Such a construction needs to be carried out in a measurable way, which motivates the following lemma. In the sequel, all optimal couplings and optimal costs considered use the cost function  $C(v, u) = |v - u|^2$ .

**Lemma 5** (coupling). *Fix  $L > 0$  and  $i \in \{1, \dots, N\}$ . Then, there exists an  $\mathbb{R}^3$ -valued function  $\Pi_t^{i,L}(\mathbf{x}, \xi)$ , measurable in  $(t, \mathbf{x}, \xi) \in [0, \infty) \times (\mathbb{R}^3)^N \times A^i$ , with the following property: for any  $(t, \mathbf{x}) \in [0, \infty) \times (\mathbb{R}^3)^N$  and any random variable  $\xi$  uniformly chosen in  $A^i$ , the pair  $(x^{i(\xi)}, \Pi_t^{i,L}(\mathbf{x}, \xi))$  is an optimal coupling between  $\bar{\mathbf{x}}^i$  and  $f_t^L$ . Moreover, for any exchangeable random vector  $\mathbf{X} \in (\mathbb{R}^3)^N$  and any bounded measurable function  $h$ , we have  $\mathbb{E} \int_{j=1}^j h(\Pi_t^{i,L}(\mathbf{X}, \xi)) d\xi = \int_{\mathbb{R}^3} h(u) f_t^L(du)$  for any  $j \in \{1, \dots, N\}$ ,  $j \neq i$ .*

*Proof.* See Lemma 3 in [5].  $\square$

To ensure that the post-collisional velocities of  $V_t^i$  and  $U_t^{i,L}$  do not differ much, we will use the functions  $\Pi^{i,L}$  of Lemma 5 to define our system  $\mathbf{U}^L = (U^{1,L}, \dots, U^{N,L})$  of cutoff nonlinear processes. This will mean that at each jump of  $v = V_t^i$  together with some other particle  $v_*$ , the corresponding process  $u = U_t^{i,L}$  will sample some  $f_t^L$ -distributed variable  $u_*$  to interact with, in such a way that the interactions of the system  $\mathbf{U}^L$  mimic those of the particle system  $\mathbf{V}$ .

However, post-collisional velocities will also depend on the angles  $\phi$  chosen in circles of the form  $\mathcal{C}(v, v_*, \theta)$  and  $\mathcal{C}(u, u_*, \theta)$  associated with each collision. As remarked by Tanaka [28,29], no continuity assumption can be made about the functions  $\hat{i}$  and  $\hat{j}$ , and, in order to control the distance between  $V_t^i$  and  $U_t^{i,L}$  after a collision, one has to make specific (non trivial) uniformly random choices for those angles as well. In the present paper, we will choose the angles  $\phi$  uniformly in the circles  $\mathcal{C}(v, v_*, \theta)$  and  $\mathcal{C}(u, u_*, \theta)$  in such a way that their joint distribution is an optimal coupling of the uniform laws on these circles, with respect to the quadratic cost. The optimal transport cost happens to have a nice explicit formula, with the optimal transport map depending only on  $v - v_*$  and  $u - u_*$ , in a fully explicit way. This is stated in the following:

**Lemma 6** (optimal coupling of circles). *Recall that  $\hat{\mathcal{C}}(b, r, d)$  denotes the circle centered at  $b \in \mathbb{R}^3$ , with radius  $r > 0$  and orthogonal direction  $d \in \mathbb{S}^2$ ; alternatively, for  $v, v_* \in \mathbb{R}^3$  and  $\theta \in [0, \pi/2]$ ,  $\mathcal{C}(v, v_*, \theta)$  denotes the circle centered at  $\frac{v+v_*}{2} + \cos \theta \frac{v-v_*}{2}$ , with radius  $\sin \theta \frac{|v-v_*|}{2}$  and orthogonal to  $\frac{v-v_*}{|v-v_*|}$ . Then:*

(i) For any  $b, \tilde{b} \in \mathbb{R}^3$ ,  $r, \tilde{r} \geq 0$  and  $d, \tilde{d} \in \mathbb{S}^2$ , the optimal transport cost between the uniform distributions on the circles  $\hat{\mathcal{C}}(b, r, d)$  and  $\hat{\mathcal{C}}(\tilde{b}, \tilde{r}, \tilde{d})$  is given by

$$\mathcal{W}_2^2 \left( \text{unif}_{\hat{\mathcal{C}}(b,r,d)}, \text{unif}_{\hat{\mathcal{C}}(\tilde{b},\tilde{r},\tilde{d})} \right) = |b - \tilde{b}|^2 + (r - \tilde{r})^2 + r\tilde{r}(1 - |d \cdot \tilde{d}|). \quad (16)$$

(ii) There exists a measurable function  $\varphi : \mathbb{R}^3 \times \mathbb{R}^3 \times [0, 2\pi) \rightarrow [0, 2\pi)$  with the following property: for every  $v, v_*, u, u_* \in \mathbb{R}^3$ ,  $\theta, \vartheta \in [0, \pi/2]$  and for any random variable  $\phi$  uniformly chosen in  $[0, 2\pi)$ , the pair

$$(v + a(v, v_*, \theta, \phi), u + a(u, u_*, \vartheta, \phi))$$

where  $\varphi = \varphi(v - v_*, u - u_*, \phi)$ , is an optimal coupling of the uniform distributions on the circles  $\mathcal{C}(v, v_*, \theta)$  and  $\mathcal{C}(u, u_*, \vartheta)$ .

*Proof.* We first prove (i). Without loss of generality, assume  $\tilde{b} = 0$  and  $d \cdot \tilde{d} \geq 0$ . Let  $h = h(d, \tilde{d}) \in \mathbb{S}^2$  be a fixed measurable choice of a unitary vector orthogonal to both  $d$  and  $\tilde{d}$  (if they are parallel then there are infinitely many such  $h$ 's; if not, there are only 2; we can take for instance  $h = \hat{v}(d)/|d|$  in the first case and  $h = d \times \tilde{d}/|d \times \tilde{d}|$  in the second). Let also  $k, \tilde{k} \in \mathbb{S}^2$  be such that  $(h, k, d)$  and  $(h, \tilde{k}, \tilde{d})$  are orthonormal bases of  $\mathbb{R}^3$  with the same orientation  $s = d \cdot (h \times k) = \tilde{d} \cdot (h \times \tilde{k}) \in \{-1, 1\}$ , so that  $k \cdot \tilde{k} = sk \cdot (\tilde{d} \times h) = s\tilde{d} \cdot (h \times k) = d \cdot \tilde{d}$ . With these bases, we can now parametrize the circles  $\hat{\mathcal{C}}(b, r, d)$  and  $\hat{\mathcal{C}}(\tilde{b}, \tilde{r}, \tilde{d})$  using angles  $\phi$  and  $\tilde{\phi} \in [0, 2\pi)$ . Namely, a point  $x \in \hat{\mathcal{C}}(b, r, d)$  is written as  $x = b + r(\cos \phi)k + r(\sin \phi)h$ , while a point  $y \in \hat{\mathcal{C}}(\tilde{b}, \tilde{r}, \tilde{d})$  is written as  $y = \tilde{r}(\cos \tilde{\phi})\tilde{k} + \tilde{r}(\sin \tilde{\phi})h$ . Then, the associated cost is

$$\begin{aligned} C(\phi, \tilde{\phi}) &= |b + r(\cos \phi)k - \tilde{r}(\cos \tilde{\phi})\tilde{k} + (r \sin \phi - \tilde{r} \sin \tilde{\phi})h|^2 \\ &= |b|^2 + r^2 + \tilde{r}^2 - 2r\tilde{r}[(\sin \phi \sin \tilde{\phi}) + (\cos \phi \cos \tilde{\phi})d \cdot \tilde{d}] \\ &\quad + 2b \cdot [r(\cos \phi)k - \tilde{r}(\cos \tilde{\phi})\tilde{k} + (r \sin \phi - \tilde{r} \sin \tilde{\phi})h]. \end{aligned}$$

Using the inequality  $2\alpha\beta \leq \alpha^2 + \beta^2$  in the cross-terms, we obtain  $C(\phi, \tilde{\phi}) \geq \Phi(\phi) - \Psi(\tilde{\phi})$  for all  $\phi, \tilde{\phi} \in [0, 2\pi)$ , where

$$\begin{aligned} \Phi(\phi) &= |b|^2 + r^2 - r\tilde{r}[(\sin \phi)^2 + (\cos \phi)^2 d \cdot \tilde{d}] + 2rb \cdot [(\cos \phi)k + (\sin \phi)h] \\ \Psi(\tilde{\phi}) &= -\{\tilde{r}^2 - r\tilde{r}[(\sin \tilde{\phi})^2 + (\cos \tilde{\phi})^2 d \cdot \tilde{d}] - 2\tilde{r}b \cdot [(\cos \tilde{\phi})\tilde{k} + (\sin \tilde{\phi})h]\}. \end{aligned}$$

Moreover, the equality  $C(\phi, \tilde{\phi}) = \Phi(\phi) - \Psi(\tilde{\phi})$  is attained when  $\phi = \tilde{\phi}$ . Using for instance Remark 5.13 of [32], this shows that taking  $\phi = \tilde{\phi}$  uniformly distributed on  $[0, 2\pi)$  in fact provides an optimal coupling of the uniform distributions on  $\mathcal{C}(b, r, d)$  and  $\mathcal{C}(\tilde{b}, \tilde{r}, \tilde{d})$ . This proves point (i), since the cost of this coupling is

$$\begin{aligned} \int_0^{2\pi} C(\phi, \phi) \frac{d\phi}{2\pi} &= \frac{1}{2\pi} \int_0^{2\pi} \{|b|^2 + r^2 + \tilde{r}^2 - 2r\tilde{r}[(\sin \phi)^2 + (\cos \phi)^2 d \cdot \tilde{d}] \\ &\quad + 2b \cdot [(\cos \phi)(rk - \tilde{r}\tilde{k}) + (\sin \phi)(r - \tilde{r})h]\} d\phi \\ &= |b|^2 + r^2 + \tilde{r}^2 - r\tilde{r}(1 + d \cdot \tilde{d}). \end{aligned}$$

We now prove (ii). Put  $d = \frac{v-v_*}{|v-v_*|}$  and  $\tilde{d} = \frac{u-u_*}{|u-u_*|}$ . For some fixed measurable choice  $(d, \tilde{d}) \mapsto h = h(d, \tilde{d})$  of a vector  $h$  orthogonal to both  $d$  and  $\tilde{d}$ , let  $\phi_i = \phi_i(d, \tilde{d}) \in [0, 2\pi)$ ,  $i = 1, 2$  be the unique angles such that

$$\frac{\Gamma(v - v_*, \phi_1)}{|v - v_*|} = \frac{\Gamma(u - u_*, \phi_2)}{|u - u_*|} =: h,$$

Note that  $(\phi_1, \phi_2)$  depend only on  $v - v_*$  and  $u - u_*$  through  $d, \tilde{d}$ ,  $\Gamma(v - v_*, \cdot)$  and  $\Gamma(u - u_*, \cdot)$ , in a measurable way. Now put

$$k := \frac{\Gamma(v - v_*, \phi_1 + \pi/2)}{|v - v_*|}, \quad \tilde{k} := \frac{\Gamma(u - u_*, \phi_2 + s\pi/2)}{|u - u_*|}.$$

Here  $s = \pm 1$  is chosen such that the rotation in  $\pi/2$  is performed with the same orientation. More specifically, if  $d \cdot \tilde{d} \geq 0$ , then  $s = 1$  when the bases  $(\frac{v-v_*}{|v-v_*|}, \frac{\hat{i}(v-v_*)}{|v-v_*|}, \frac{\hat{j}(v-v_*)}{|v-v_*|})$  and  $(\frac{u-u_*}{|u-u_*|}, \frac{\hat{i}(u-u_*)}{|u-u_*|}, \frac{\hat{j}(u-u_*)}{|u-u_*|})$  have the same orientation, and  $s = -1$  otherwise; but when  $d \cdot \tilde{d} < 0$ , we make the opposite choice. Now, the same argument of part (i) shows that if  $\phi$  is a uniform random variable on  $[0, 2\pi)$  then  $v + a(v, v_*, \theta, \phi)$  and  $u + a(u, u_*, \vartheta, s(\phi - \phi_1) + \phi_2)$  constitute an optimal coupling. Put  $\varphi = s(\phi - \phi_1) + \phi_2$  and the conclusion follows.  $\square$

*Remark 7.*

- The expression on the right of (16) is nice: the term  $|b - \tilde{b}|^2$  is the cost associated to translation of the circles, the term  $(r - \tilde{r})^2$  is the dilation or contraction cost, and  $r\tilde{r}(1 - |d \cdot \tilde{d}|)$  corresponds to inclination.
- When  $(v - v_*) \cdot (u - u_*) \geq 0$  and  $\theta = \vartheta$ , the coupling given in Lemma 6-(ii) reduces to the parallel spherical coupling of [25].

With the functions  $\Pi^{i,L}$  and  $\varphi$  of Lemmas 5 and 6-(ii) in hand, we can now introduce, at a formal level first, a system of cutoff nonlinear processes  $\mathbf{U}^L = (U^{1,L}, \dots, U^{N,L})$ , suitably constructed in the same probability space as  $\mathbf{V} = (V^1, \dots, V^N)$  defined in (13). Recall that the pair  $(\mathbf{V}_0, \mathbf{U}_0)$  is given and specified in (12). Mimicking (15), for each  $L \in [1, \infty)$ ,  $N \in \mathbb{N}$  and  $i = 1, \dots, N$ , the processes  $U^{i,L}$  is defined as the solution, starting from  $U_0^i$ , of the stochastic equation

$$dU_t^{i,L} = \int_0^\infty \int_0^{2\pi} \int_0^N c_L(U_{t-}^{i,L}, \Pi_t^{i,L}(\mathbf{U}_{t-}^L, \xi), z, \varphi_{t-}^i) \mathcal{N}^i(dt, dz, d\phi, d\xi), \quad (17)$$

where we have used the shorthand

$$\varphi_{t-}^i = \varphi(V_{t-}^i - V_{t-}^{i(\xi)}, U_{t-}^{i,L} - \Pi_t^{i,L}(\mathbf{U}_{t-}^L, \xi), \phi). \quad (18)$$

In words, at jump instants of  $U_t^{i,L}$ , this process collides with an  $f_t^L$ -distributed random variable, which is optimally coupled to the realization  $U_{t-}^{i(\xi),L}$  of the (random) measure  $\tilde{\mathbf{U}}_{t-}^{i,L}$ . Since the Poisson measures  $\mathcal{N}^i$  and  $\mathcal{N}^j$  share some of its atoms, processes  $U^{i,L}$  and  $U^{j,L}$  have simultaneous jumps and hence are not independent.

We can write the following joint SDE for the pair  $(\mathbf{V}, \mathbf{U}^L)$ , arranged as the collection of pairs  $((V^1, U^{1,L}), \dots, (V^N, U^{N,L})) \in (\mathbb{R}^3 \times \mathbb{R}^3)^N$ :

$$d(\mathbf{V}, \mathbf{U}^L)_t = \int_0^\infty \int_0^{2\pi} \int_{[0,N]^2} \sum_{i \neq j} \mathbf{1}_{\{i(\xi)=i, i(\zeta)=j\}} \times \mathbf{b}_{L,ij}(\mathbf{V}_{t^-}, \mathbf{U}_{t^-}^L, t, z, \phi, \xi, \zeta) \mathcal{N}(dt, dz, d\phi, d\xi, d\zeta), \tag{19}$$

where  $\mathbf{b}_{L,ij}(\mathbf{x}, \mathbf{y}, t, z, \phi, \xi, \zeta) \in (\mathbb{R}^3 \times \mathbb{R}^3)^N$  is given by

$$\begin{aligned} & (\mathbf{b}_{L,ij}(\mathbf{x}, \mathbf{y}, t, z, \phi, \xi, \zeta))^\ell \\ &= \begin{cases} \left( c(x^i, x^j, z, \phi), c_L(y^i, \Pi_t^{i,L}(\mathbf{y}, \zeta), z, \varphi(x^i - x^j, y^i - \Pi_t^{i,L}(\mathbf{y}, \zeta), \phi)) \right) & \text{if } \ell = i, \\ \left( c(x^j, x^i, z, \phi), c_L(y^j, \Pi_t^{j,L}(\mathbf{y}, \xi), z, \varphi(x^j - x^i, y^j - \Pi_t^{j,L}(\mathbf{y}, \xi), \phi)) \right) & \text{if } \ell = j, \\ (0, 0) & \text{otherwise.} \end{cases} \end{aligned}$$

Existence for each  $L \in [1, \infty)$  and  $N \in \mathbb{N}$  of a pair  $(\mathbf{V}, \mathbf{U}^L)$  solving (19), along with its relevant properties, is stated in the next result. Some arguments of the proof are standard or can be adapted from previous works, so details will be provided only when needed. The proof is given in the Appendix.

**Lemma 8.** *Assume (5) and let  $L \in [1, \infty)$  and  $N \in \mathbb{N}$ . We have:*

- (i) *If  $\mathbf{V}$  is a solution to (13), then it has the law of the unique  $(\mathbb{R}^3)^N$  valued Markov process with generator given by (3). In particular, we almost surely have  $\sum_{i=1}^N V_t^i = \sum_{i=1}^N V_0^i$  and  $\sum_{i=1}^N |V_t^i|^2 = \sum_{i=1}^N |V_0^i|^2$  for all  $t \geq 0$ .*
- (ii) *There is weak existence and uniqueness of a solution  $(\mathbf{V}, \mathbf{U}^L)$  to the system of SDEs (19).*
- (iii) *For each  $i = 1, \dots, N$ , the process  $U^{i,L}$  is a cutoff nonlinear process, and in particular we have  $\mathcal{L}(U_t^{i,L}) = f_t^L$ .*
- (iv) *Last, the collection of pairs of processes  $(V^1, U^{1,L}), \dots, (V^N, U^{N,L})$  is exchangeable.*

*Remark 9.* One can also check, using the preservation of moments of the processes  $\mathbf{U}^L$ ,  $L \in [1, \infty)$  and Lemma 15, that the family of laws of  $(\mathbf{V}, \mathbf{U}^L)$  has accumulation points as  $L \rightarrow \infty$  which are couplings of the particle system (13) and a system of  $N$  non-independent nonlinear processes. Unfortunately, due to the lack of continuity of the functions  $\varphi$  and  $\Pi_t^{i,L}$ , this does not readily ensure (weak) well-posedness for the system (19) in the case  $L = \infty$ , which would simplify the construction and proofs. This is the reason why we are constrained to work with a system of cutoff nonlinear processes.

### 3. Estimates and Technical Results

We will use the following bounds a couple of times: under (5), it can be easily seen that for some constants  $0 < c_2 < c_3$  we have:

$$c_2(1+z)^{-1/v} \leq G(z) \leq c_3(1+z)^{-1/v} \quad \forall z > 0. \tag{20}$$

The following lemmas provide useful estimates for our purposes. Typically, one wants to use these lemmas with  $v$  and  $v_*$  taken from the particle system, and  $u$  and  $u_*$  taken from the system of cutoff nonlinear processes.

**Lemma 10.** Write  $R(v, u) := |v||u| + |v \cdot u| - 2v \cdot u \geq 0$ . For any  $v, v_*, u, u_* \in \mathbb{R}^3$ ,  $\theta, \vartheta \in [0, \pi/2]$ , write  $\varphi = \varphi(v - v_*, u - u_*, \phi)$ . Then:

$$\begin{aligned} & \int_0^{2\pi} \left( |v + a(v, v_*, \theta, \phi) - u - a(u, u_*, \vartheta, \varphi)|^2 - |v - u|^2 \right) \frac{d\phi}{2\pi} \\ &= - \left[ (v - u) + (v_* - u_*) \right] \cdot \left[ \frac{1 - \cos \theta}{2} (v - v_*) - \frac{1 - \cos \vartheta}{2} (u - u_*) \right] \\ & \quad - \frac{\sin \theta \sin \vartheta}{4} R(v - v_*, u - u_*) + \frac{1 - \cos(\theta - \vartheta)}{2} (v - v_*) \cdot (u - u_*). \end{aligned} \tag{21}$$

*Proof.* Setting

$$\begin{aligned} b &= \frac{v + v_*}{2} + \cos \theta \frac{v - v_*}{2}, & r &= \sin \theta \frac{|v - v_*|}{2}, & d &= \frac{v - v_*}{|v - v_*|}, \\ \tilde{b} &= \frac{u + u_*}{2} + \cos \vartheta \frac{u - u_*}{2}, & \tilde{r} &= \sin \vartheta \frac{|u - u_*|}{2}, & \tilde{d} &= \frac{u - u_*}{|u - u_*|}, \end{aligned}$$

we have

$$\begin{aligned} & 4(b - \tilde{b})^2 \\ &= |(v - u) + (v_* - u_*) + \cos \theta (v - v_*) - \cos \vartheta (u - u_*)|^2 \\ &= |v - u|^2 + |v_* - u_*|^2 + 2(v - u) \cdot (v_* - u_*) \\ & \quad + \cos^2 \theta |v - v_*|^2 + \cos^2 \vartheta |u - u_*|^2 - 2 \cos \theta \cos \vartheta (v - v_*) \cdot (u - u_*) \\ & \quad + 2[(v - u) + (v_* - u_*)] \cdot [\cos \theta (v - v_*) - \cos \vartheta (u - u_*)] \\ &= 3|v - u|^2 - |v_* - u_*|^2 + 2(v - u) \cdot (v_* - u_*) \\ & \quad + \cos^2 \theta |v - v_*|^2 + \cos^2 \vartheta |u - u_*|^2 - 2 \cos \theta \cos \vartheta (v - v_*) \cdot (u - u_*) \\ & \quad - 2[(v - u) + (v_* - u_*)] \cdot [(1 - \cos \theta)(v - v_*) - (1 - \cos \vartheta)(u - u_*)] \end{aligned} \tag{22}$$

and

$$\begin{aligned} & 4(r - \tilde{r})^2 + 4r\tilde{r}(1 - |d \cdot \tilde{d}|) \\ &= \sin^2 \theta |v - v_*|^2 + \sin^2 \vartheta |u - u_*|^2 \\ & \quad - \sin \theta \sin \vartheta (|v - v_*||u - u_*| + |(v - v_*) \cdot (u - u_*)|) \\ &= \sin^2 \theta |v - v_*|^2 + \sin^2 \vartheta |u - u_*|^2 - \sin \theta \sin \vartheta R(v - v_*, u - u_*) \\ & \quad - 2 \sin \theta \sin \vartheta (v - v_*) \cdot (u - u_*). \end{aligned} \tag{23}$$

Adding (22) and (23), using that  $|v - v_*|^2 + |u - u_*|^2 + 2(v - u) \cdot (v_* - u_*) = |v - u|^2 + |v_* - u_*|^2 + 2(v - v_*) \cdot (u - u_*)$  and the identity  $\cos \theta \cos \vartheta + \sin \theta \sin \vartheta = \cos(\theta - \vartheta)$ , yields

$$\begin{aligned} & 4(b - \tilde{b})^2 + 4(r - \tilde{r})^2 + 4r\tilde{r}(1 - |d \cdot \tilde{d}|) \\ &= 4|v - u|^2 - \sin \theta \sin \vartheta R(v - v_*, u - u_*) \\ & \quad + 2(1 - \cos(\theta - \vartheta))(v - v_*) \cdot (u - u_*) \\ & \quad - 2[(v - u) + (v_* - u_*)] \cdot [(1 - \cos \theta)(v - v_*) - (1 - \cos \vartheta)(u - u_*)]. \end{aligned}$$

Thanks to Lemma 6-(ii),  $\varphi$  is an optimal transport map, and then the integral on the left side of (21) without the term  $-|v - u|^2$  is actually the cost given by Lemma 6-(i), that is,  $(b - \tilde{b})^2 + (r - \tilde{r})^2 + r\tilde{r}(1 - |d \cdot \tilde{d}|)$ . Dividing by 4 and subtracting  $|v - u|^2$  in the above identity, the result follows.  $\square$

**Corollary 11.** *Assume (5). Fix any  $K, L \in [0, \infty]$  with  $K \geq L$ , and define  $\Phi_L^K := \int_L^K \frac{1 - \cos G(z)}{2} dz \geq 0$ . For any  $v, v_*, u, u_* \in \mathbb{R}^3$ , write  $\varphi = \varphi(v - v_*, u - u_*, \phi)$ . Then we have*

$$\begin{aligned} & \int_0^\infty \int_0^{2\pi} \left( |v + c_K(v, v_*, z, \phi) - u - c_L(u, u_*, z, \varphi)|^2 - |v - u|^2 \right) \frac{d\phi}{2\pi} dz \\ &= \Phi_0^L [-|v - u|^2 + |v_* - u_*|^2] + \Phi_L^K (v - v_*) \cdot (2u - v - v_*) \\ &\quad - R(v - v_*, u - u_*) \int_0^L \frac{\sin^2 G(z)}{4} dz. \\ &\leq \Phi_0^L [-|v - u|^2 + |v_* - u_*|^2] + C(|v| + |v_*| + |u|)^2 (1 + L)^{1-2/\nu}. \end{aligned}$$

*Proof.* Split the integral with respect to  $z$  into  $\int_0^L$  and  $\int_L^K$ . For the first integral we have  $c_K(v, v_*, z, \phi) = a(v, v_*, \theta, \phi)$  and  $c_L(u, u_*, z, \varphi) = a(u, u_*, \theta, \varphi)$  for  $\theta = G(z)$ ; using Lemma 10 yields the first and third terms in the equality. For the second integral we have  $c_L(u, u_*, z, \varphi) = a(v, v_*, 0, \varphi)$ , so this time we use Lemma 10 with  $\theta = G(z)$  and  $\vartheta = 0$ , which gives the second term. The inequality is then obtained discarding the negative third term, noting that  $\Phi_L^K \leq C \int_L^\infty G^2(z) dz$ , and using (20).  $\square$

The next lemma is of key importance, since it gives a decoupling estimate for the system of non-independent cutoff nonlinear processes  $\mathbf{U}^L$ . The proof, also relying on a coupling argument, follows Lemma 6 of [5].

**Lemma 12** (decoupling). *Assume (5) and take  $L \in [1, \infty)$ . Then, there exists a constant  $C$  independent of  $L$  such that for all  $k \in \{1, \dots, N\}$  and all  $t \geq 0$ ,*

$$\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{U}_t^L), (f_t^L)^{\otimes k}) \leq C \frac{k}{N}.$$

*Proof.* Given  $k \in \{1, \dots, N\}$  fixed, we will construct  $k$  independent cutoff nonlinear processes  $\tilde{U}^{1,L}, \dots, \tilde{U}^{k,L}$  such that  $\mathbb{E}|U_t^{i,L} - \tilde{U}_t^{i,L}|^2$  is small, for all  $i = 1, \dots, k$ . To achieve this, the idea is the following: when  $U_t^{i,L}$  has a simultaneous jump with some  $U_t^{j,L}$  with  $j \notin \{1, \dots, k\}$ , then the process  $\tilde{U}_t^{i,L}$  will use the same sample of  $f_t^L$  used by  $U_t^{j,L}$  to define its own jump; but when  $j \in \{1, \dots, k\}$ , then one of the processes,  $\tilde{U}_t^{i,L}$  or  $\tilde{U}_t^{j,L}$ , will not jump at that instant. We will then use an additional, independent source of randomness to define new jumps that compensate for the missing ones. Since, when  $k \ll N$ , the second kind of jump occurs much less frequently, this construction will give the desired estimate.

Consider a Poisson point measure  $\mathcal{M}$  that is an independent copy of  $\mathcal{N}$ , also independent from  $(\mathbf{V}_0, \mathbf{U}_0)$ , and define for each  $i \in \{1, \dots, k\}$

$$\begin{aligned} \mathcal{M}^i(dt, dz, d\phi, d\xi) &= \mathcal{N}(dt, dz, d\phi, [i - 1, i], d\xi) \\ &\quad + \mathcal{N}(dt, dz, d\phi, d\xi, [i - 1, i]) \mathbf{1}_{[k, N]}(\xi) \\ &\quad + \mathcal{M}(dt, dz, d\phi, d\xi, [i - 1, i]) \mathbf{1}_{[0, k]}(\xi). \end{aligned}$$



That is,  $\mathcal{M}^i$  selects the atoms of  $\mathcal{N}(dt, dz, d\phi, d\xi, d\zeta)$  where either  $(\mathbf{i}(\xi) = i)$ , or  $(\mathbf{i}(\zeta) = i$  and  $\mathbf{i}(\xi) \notin \{1, \dots, k\})$ , and to make up for the dropped atoms it also selects new ones from  $\mathcal{M}(dt, dz, d\phi, d\xi, d\zeta)$ , where  $\mathbf{i}(\zeta) = i$  and  $\mathbf{i}(\xi) \in \{1, \dots, k\}$ . This ensures that no such atom appears in two  $\mathcal{M}^i$ 's, implying that they are independent Poisson point measures, all with intensity  $dt dz d\phi d\xi \mathbf{1}_{A^i}(\xi)/[2(N - 1)\pi]$ , just like  $\mathcal{N}^i$ .

Mimicking (17), we define  $\tilde{U}^{i,L}$  as the solution, starting from  $\tilde{U}_0^{i,L} = U_0^i$ , of the stochastic equation

$$d\tilde{U}_t^{i,L} = \int_0^\infty \int_0^{2\pi} \int_0^N c_L(\tilde{U}_t^{i,L}, \Pi_t^{i,L}(\mathbf{U}_t^L, \xi), z, \tilde{\varphi}_t^{i,-}) \mathcal{M}^i(dt, dz, d\phi, d\xi). \tag{24}$$

Here we write

$$\tilde{\varphi}_t^{i,-} = \varphi(U_t^{i,L} - \Pi_t^{i,L}(\mathbf{U}_t^L, \xi), \tilde{U}_t^{i,L} - \Pi_t^{i,L}(\mathbf{U}_t^L, \xi), \varphi_t^{i,-}),$$

where  $\varphi_t^{i,-}$  was defined in (18). In other words:  $\tilde{\varphi}_t^{i,-}$  takes the angle  $\varphi_t^{i,-}$  and maps it to  $[0, 2\pi)$  in such a way that the resulting pair  $(\varphi_t^{i,-}, \tilde{\varphi}_t^{i,-})$  parametrizes (as a function of  $\phi$ ) an optimal coupling of the uniform distributions on the circles with orthogonal directions  $U_t^{i,L} - \Pi_t^{i,L}(\mathbf{U}_t^L, \xi)$  and  $\tilde{U}_t^{i,L} - \Pi_t^{i,L}(\mathbf{U}_t^L, \xi)$ , whenever  $\phi$  is uniformly chosen on  $[0, 2\pi)$ .

The latter ensures closeness of the states of  $U_t^{i,L}$  and  $\tilde{U}_t^{i,L}$  after the joint jump.

If we define  $\tilde{\mathcal{M}}^{i,L}(dt, dz, d\phi, dv)$  to be the point measure on  $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times \mathbb{R}^3$  with atoms  $(t, z, \tilde{\varphi}_t^{i,-}, \Pi_t^{i,L}(\mathbf{U}_t^L, \xi))$  for every atom  $(t, z, \phi, \xi)$  of  $\mathcal{M}^i$ , it is clear that  $\tilde{U}^{i,L}$  depends only on  $\tilde{\mathcal{M}}^{i,L}$  and  $U_0^i$ . Since: (1) the dependence on  $\mathbf{V}$  and  $\mathbf{U}^L$  in (24) is predictable with respect to  $\mathcal{N}$ ,  $\mathcal{M}$  and the initial data, (2) the Poisson measures  $\mathcal{M}^1, \dots, \mathcal{M}^k$  are independent, (3) the  $\xi$ -law of  $\Pi_t^{i,L}(\mathbf{x}, \xi)$  is  $f_t^L$  for every  $\mathbf{x} \in \mathbb{R}^N$ , and (4) the  $\phi$ -law of  $\varphi(v, u, \phi)$  is the uniform distribution on  $[0, 2\pi)$  for any  $v, u \in \mathbb{R}^3$ , one can use the compensation formula to compute the joint Laplace functional of  $\tilde{\mathcal{M}}^{1,L}, \dots, \tilde{\mathcal{M}}^{k,L}$  and deduce that they are independent Poisson point measures, all with intensity  $dt dz d\phi f_t^L(dv)/2\pi$ . This shows that  $\tilde{U}^{i,L}$  satisfies (11) with  $\mathcal{M}^L$  replaced by  $\tilde{\mathcal{M}}^{i,L}$ , and then  $\tilde{U}^{1,L}, \dots, \tilde{U}^{k,L}$  are independent cutoff nonlinear processes.

Consequently, we have

$$\mathcal{W}_2^2(\mathcal{L}^k(\mathbf{U}_t^L), (f_t^L)^{\otimes k}) \leq \mathbb{E} \frac{1}{k} \sum_{j=1}^k |U_t^{j,L} - \tilde{U}_t^{j,L}|^2 = \mathbb{E} |U_t^{i,L} - \tilde{U}_t^{i,L}|^2.$$

Thus, it suffices to estimate the quantity  $h_t := \mathbb{E}|U_t^{i,L} - \tilde{U}_t^{i,L}|^2$ , for any  $i \in \{1, \dots, k\}$  fixed. We can write

$$h_t - h_s = J_{s,t}^1 + J_{s,t}^2 + J_{s,t}^3 \tag{25}$$

where  $J_{s,t}^1$  is the term associated with simultaneous jumps of  $U^{i,L}$  and  $\tilde{U}^{i,L}$ ,  $J_{s,t}^2$  corresponds to jumps of  $U^{i,L}$  alone, and  $J_{s,t}^3$  corresponds to jumps of  $\tilde{U}^{i,L}$  alone. To write this terms explicitly, let us first shorten notation: write  $U_r := U_r^{i,L}$ ,  $\tilde{U}_r := \tilde{U}_r^{i,L}$ ,  $\Pi_r := \Pi_r^{i,L}(\mathbf{U}_r^L, \xi)$ ,  $c_r := c(U_r, \Pi_r, z, \varphi_r^i)$  and  $\tilde{c}_r := c(\tilde{U}_r, \Pi_r, z, \tilde{\varphi}_r^i)$ . From (17) and (24),  $J_{s,t}^1$ ,  $J_{s,t}^2$  and  $J_{s,t}^3$  are thus given by

$$\begin{aligned}
 J_{s,t}^1 &= \mathbb{E} \int_{(s,t]} \int_0^L \int_0^{2\pi} \int_0^N \left( |U_{r-} + c_{r-} - \tilde{U}_{r-} - \tilde{c}_{r-}|^2 - |U_{r-} - \tilde{U}_{r-}|^2 \right) \\
 &\quad \left[ \mathcal{N}(dr, dz, d\phi, [i-1, i), d\xi) + \mathcal{N}(dr, dz, d\phi, d\xi, [i-1, i)) \mathbf{1}_{[k, N)}(\xi) \right], \\
 J_{s,t}^2 &= \mathbb{E} \int_{(s,t]} \int_0^L \int_0^{2\pi} \int_0^N \left( |U_{r-} + c_{r-} - \tilde{U}_{r-}|^2 - |U_{r-} - \tilde{U}_{r-}|^2 \right) \\
 &\quad \mathcal{N}(dr, dz, d\phi, d\xi, [i-1, i)) \mathbf{1}_{[0, k)}(\xi), \\
 J_{s,t}^3 &= \mathbb{E} \int_{(s,t]} \int_0^L \int_0^{2\pi} \int_0^N \left( |U_{r-} - \tilde{U}_{r-} - \tilde{c}_{r-}|^2 - |U_{r-} - \tilde{U}_{r-}|^2 \right) \\
 &\quad \mathcal{M}(dr, dz, d\phi, d\xi, [i-1, i)) \mathbf{1}_{[0, k)}(\xi).
 \end{aligned}$$

Recall that  $\mathcal{N}$  and  $\mathcal{M}$  have intensity  $dt dz d\phi d\xi d\zeta \mathbf{1}_{\mathcal{G}}(\xi, \zeta) / [4(N-1)\pi]$ . Note that  $\int_{i-1}^i \mathbf{1}_{\mathcal{G}}(\xi, \zeta) d\zeta = \mathbf{1}_{A^i}(\xi)$ , where  $A^i = [0, N) \setminus [i-1, i)$ . Using the compensation formula, the Poisson point measures in the integrals can be replaced by their intensities, and we obtain for  $J_{s,t}^1$ :

$$\begin{aligned}
 J_{s,t}^1 &= \mathbb{E} \int_s^t \int_0^L \int_0^{2\pi} \int_0^N \left( |U_r + c_r - \tilde{U}_r - \tilde{c}_r|^2 \right. \\
 &\quad \left. - |U_r - \tilde{U}_r|^2 \right) [\mathbf{1}_{A^i}(\xi) + \mathbf{1}_{[k, N)}(\xi)] \frac{dr dz d\phi d\xi}{4(N-1)\pi} \\
 &\leq -\Phi_0^L \mathbb{E} \int_s^t \int_0^N |U_r - \tilde{U}_r|^2 [\mathbf{1}_{A^i}(\xi) + \mathbf{1}_{[k, N)}(\xi)] \frac{dr d\xi}{2(N-1)} \\
 &\leq -\frac{\Phi_0^L}{2} \int_s^t h_r dr,
 \end{aligned} \tag{26}$$

where we have used Corollary 11 with  $v = U_r, u = \tilde{U}_r, v_* = u_* = \Pi_r, K = L$  and the change of variable  $(\phi, \varphi_r^i) \mapsto (\phi_r^i, \tilde{\varphi}_r^i)$ . For  $J_{s,t}^2$  we get:

$$\begin{aligned}
 J_{s,t}^2 &= \mathbb{E} \int_s^t \int_0^L \int_0^{2\pi} \int_0^N \left( |U_r + c_r - \tilde{U}_r|^2 - |U_r - \tilde{U}_r|^2 \right) \mathbf{1}_{A^i \cap [0, k)}(\xi) \frac{dr dz d\phi d\xi}{4(N-1)\pi} \\
 &\leq C \mathbb{E} \int_s^t \int_0^N (|U_r| + |\Pi_r| + |\tilde{U}_r|)^2 \mathbf{1}_{A^i \cap [0, k)}(\xi) \frac{dr d\xi}{2(N-1)} \\
 &\leq C(t-s) \frac{k-1}{N-1},
 \end{aligned} \tag{27}$$

where in the second step we have used Corollary 11 again, with  $v = U_r, u = \tilde{U}_r, v_* = u_* = \Pi_r$  and the roles of  $K$  and  $L$  exchanged (the smallest one being equal to 0). In the last step we have used the fact that  $f_r^L$  has uniformly bounded moments of order 2, that  $\mathcal{L}(U_r) = \mathcal{L}(\tilde{U}_r) = f_r^L$ , and that  $\mathbb{E} \int_{j-1}^j |\Pi_r|^2 d\xi = \int_{\mathbb{R}^3} |u|^2 f_r^L(du)$  for all  $j \neq i$ , thanks to Lemma 5. Similarly for  $J_{s,t}^3$ : using Corollary 11 with  $v = \tilde{U}_r, u = U_r, v_* = u_* = \Pi_r$ , the bound (27) is also valid for  $J_{s,t}^3$ .

Thus, from (25), (26) and (27), we obtain that  $\partial_t h_t \leq -(\Phi_0^L/2)h_t + Ck/N$  for almost all  $t \geq 0$ , and, since  $h_0 = 0$ , the conclusion follows from Gronwall's lemma (the dependence on  $L$  can be dropped since  $\Phi_0^L$  is bounded away from 0 when  $L \geq 1$  thanks to the lower bound in (20)).  $\square$

**4. Proof of Theorem 1**

For a probability measure  $\mu$  on  $\mathbb{R}^d$ , call  $\varepsilon_n(\mu) := \mathbb{E}\mathcal{W}_2^2(\mu, \bar{\mathbf{Z}})$ , where  $\mathbf{Z} = (Z^1, \dots, Z^n) \in (\mathbb{R}^d)^n$  is a vector of  $n$  independent and  $\mu$ -distributed random variables on  $\mathbb{R}^d$ . The best general estimate available for  $\varepsilon_n(\mu)$  is the following, whose proof can be found in [11]: for any  $p > 4$ , there exists a constant  $C_p < \infty$  such that for every  $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$\varepsilon_n(\mu) \leq \frac{C_p(\int |v|^p \mu(dv))^{2/p}}{n^{1/2}}. \tag{28}$$

The following lemma will allow us to work with  $\mathcal{W}_2^2(\mathcal{L}^n(\mathbf{U}_t^L), (f_t^L)^{\otimes n})$  instead of  $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^{i,L}, f_t^L)$ , but at the price of the extra term  $\varepsilon_n(f_t^L)$ :

**Lemma 13.** *Let  $\mathbf{X} = (X^1, \dots, X^m) \in (\mathbb{R}^d)^m$  be an exchangeable random vector, and let  $\mu \in \mathcal{P}(\mathbb{R}^d)$ . Then, for any  $n \leq m$ ,*

$$\begin{aligned} \frac{1}{2}\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{X}}, \mu) &\leq \frac{kn}{m} \left( \mathcal{W}_2^2(\mathcal{L}^n(\mathbf{X}), \mu^{\otimes n}) + \varepsilon_n(\mu) \right) \\ &\quad + \frac{\ell}{m} \left( \mathcal{W}_2^2(\mathcal{L}^\ell(\mathbf{X}), \mu^{\otimes \ell}) + \varepsilon_\ell(\mu) \right), \end{aligned}$$

where  $k$  and  $\ell$  are the unique non-negative integers satisfying  $m = kn + \ell$  and  $\ell \leq n - 1$ .

*Proof.* See the proof of Lemma 7 of [5].  $\square$

**Corollary 14.** *Assume (5) and (6), and take  $L \in [1, \infty)$ . Then, there exists a constant  $C > 0$  independent of  $L$ , such that for any  $i \in \{1, \dots, N\}$  and for all  $t \geq 0$ ,*

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^{i,L}, f_t^L) \leq CN^{-1/3}.$$

Moreover, the same bound is valid with  $\bar{\mathbf{U}}_t^L$  in place of  $\bar{\mathbf{U}}_t^{i,L}$ .

*Proof.* In the notation of Lemma 13, set  $m = N - 1$ ,  $\mathbf{X} = (U_t^{j,L})_{j \neq i}$  and  $\mu = f_t^L$  and given  $n \leq m$ , simply bound  $\frac{kn}{m} \leq 1$ ,  $\frac{\ell}{m} \leq \frac{n}{N}$ ,  $\mathcal{W}_2^2(\mathcal{L}^\ell(\mathbf{X}), \mu^{\otimes \ell}) \leq 4E$  and  $\varepsilon_\ell(\mu) \leq 4E$ , where  $E = \int |v|^2 f_t^L(dv) = \int |v|^2 f_0(dv)$ . Using that result we get

$$\begin{aligned} \frac{1}{2}\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_t^{i,L}, f_t^L) &\leq \mathcal{W}_2^2(\mathcal{L}^n(\mathbf{U}_t^L), (f_t^L)^{\otimes n}) + \varepsilon_n(f_t^L) + 8E \frac{n}{N} \\ &\leq C \frac{n}{N} + C \frac{1}{n^{1/2}}, \end{aligned}$$

where we have used Lemma 12 together with (28) with  $p = p_0 > 4$  and the uniformity of the  $p_0$ -moments of  $f_t^L$ . Choosing  $n = \lfloor N^{2/3} \rfloor$  yields the desired result. To obtain the same estimate with  $\bar{\mathbf{U}}_t^L$  on the left hand side, use  $m = N$  and  $\mathbf{X} = \mathbf{U}_t^L$ .  $\square$

We need to make sure that the cutoff  $L$  can be removed in a satisfactory manner; for instance, we can use Theorem 5.2 of [12]. However, for the reader’s convenience, we state here a result specific for the cutoff we use, with a shorter proof and better dependence on time:

**Lemma 15.** *Assume (5) and (6). Then there exists a constant  $C$  such that for all  $t \geq 0$  and all  $L > 0$ ,*

$$\mathcal{W}_2^2(f_t^L, f_t) \leq CtL^{1-2/\nu}.$$

*Proof.* If  $f_0$  is a Dirac mass, then  $f_t = f_t^L = f_0$  for all  $t \geq 0$  and the result is trivial. If  $f_0$  is not a Dirac mass, we know that  $f_t$  has a density for  $t > 0$  thanks to Theorem 4; therefore, there exists an optimal transport map  $T_t^L : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  such that for any random vector  $X \in \mathbb{R}^3$  with law  $f_t$ , the pair  $(X, T_t^L(X))$  is an optimal coupling between  $f_t$  and  $f_t^L$ . Moreover, thanks to the measurability of the flows  $t \mapsto (f_t, f_t^L)$  and to Theorem 1.1 in [8], the maps  $T_t^L$  can be chosen in such a way that the mapping  $(t, v) \mapsto (v, T_t^L(v))$  is measurable. Now, given a (weak) solution  $(W, \mathcal{M})$  to (10), define a process  $W^L$  as the unique, jump-by-jump solution, starting from  $W_0^L = W_0$ , to the stochastic equation

$$dW_t^L = \int_0^\infty \int_0^{2\pi} \int_{\mathbb{R}^3} c_L(W_{t-}^L, T_t^L(v), z, \varphi_{t-}) \mathcal{M}(dt, dz, d\phi, dv),$$

with  $\varphi_{t-} = \varphi(W_{t-} - v, W_{t-}^L - T_t^L(v), \phi)$ . Arguing as in the proof of part (iii) of Lemma 8, one can verify that  $W^L$  is a cutoff nonlinear process, and in particular,  $W_t^L$  has law  $f_t^L$  for each  $t \geq 0$ . For  $h_t = \mathbb{E}|W_t - W_t^L|^2$ , we obtain from this and (10):

$$\begin{aligned} \partial_t h_t &= \mathbb{E} \int_0^\infty \int_0^{2\pi} \int_{\mathbb{R}^3} \left( |W_t + c(W_t, v, z, \phi) - W_t^L - c_L(W_t^L, T_t^L(v), z, \varphi_t)|^2 \right. \\ &\quad \left. - |W_t - W_t^L|^2 \right) \frac{dtd\phi f_t(dv)}{2\pi} \\ &\leq \mathbb{E} \int_{\mathbb{R}^3} \left( -\Phi_0^L |W_t - W_t^L|^2 + \Phi_0^L |v - T_t^L(v)|^2 \right. \\ &\quad \left. + C(|W_t| + |v| + |W_t^L|)^2 L^{1-2/\nu} \right) f_t(dv), \end{aligned}$$

where we have used Corollary 11 with  $K = \infty$ . By construction, we have  $\int_{\mathbb{R}^3} |v - T_t^L(v)|^2 f_t(dv) = \mathcal{W}_2^2(f_t^L, f_t) \leq h_t$ , and using the preservation of the second moment for  $f_t^L$  and  $f_t$ , the last inequality yields  $\partial_t h_t \leq CL^{1-2/\nu}$ . Since  $h_0 = 0$ , the result follows.  $\square$

We are now ready to prove Theorem 1.

*Proof of Theorem 1.* Take  $L \in [1, \infty)$ . For some  $i \in \{1, \dots, N\}$  fixed, we will estimate the quantity  $h_t := \mathbb{E}|V_t^i - U_t^{i,L}|^2$ . To shorten notation, call  $V_r := V_r^i, V_r^i := V_r^{i(\xi)}, U_r := U_r^{i,L}$  and  $\Pi_r := \Pi_r^{i,L}(U_r^L, \xi)$ . From (15) and (17) it follows that for every  $0 \leq s < t$ ,

$$\begin{aligned} h_t - h_s &= \mathbb{E} \int_{(s,t]} \int_0^\infty \int_0^{2\pi} \int_0^N \left( \left| V_{r-} + c(V_{r-}, V_{r-}^i, z, \phi) - U_{r-} \right. \right. \\ &\quad \left. \left. - c_L(U_{r-}, \Pi_{r-}, z, \varphi_{r-}^i) \right|^2 - |V_{r-} - U_{r-}|^2 \right) \mathcal{N}^i(dr, dz, d\phi, d\xi). \end{aligned}$$

Using the compensation formula,  $\mathcal{N}^i$  can be replaced by its intensity  $drdzd\phi d\xi \mathbf{1}_{A^i}(\xi) / [2(N-1)\pi]$ , where  $A^i = [0, N] \setminus [i-1, i]$ . Corollary 11 with  $v = V_r, v_* = V_r^i, u = U_r, u_* = \Pi_r$  and  $K = \infty$  yields

$$\begin{aligned} h_t - h_s &\leq \mathbb{E} \int_s^t \int_{A^i} \left[ -\Phi_0^L |V_r - U_r|^2 + \Phi_0^L |V_r^i - \Pi_r|^2 \right. \\ &\quad \left. + C(|V_r| + |V_r^i| + |U_r|)^2 L^{1-2/\nu} \right] \frac{drd\xi}{N-1}. \end{aligned}$$

Note that  $|V_r^i - \Pi_r|^2 \leq |V_r^i - U_r^i|^2 + 2|V_r^i - U_r^i||U_r^i - \Pi_r| + |U_r^i - \Pi_r|^2$ , where  $U_r^i := U_r^{i(\xi),L}$ , and that  $\mathbb{E} \int_{A^i} |V_r^i - U_r^i|^2 d\xi / (N - 1) = h_r$  by exchangeability. Also, thanks to Lemma 5, we know that  $\int_{A^i} |U_r^i - \Pi_r|^2 d\xi / (N - 1) = \mathcal{W}_2^2(\bar{U}_r^{i,L}, f_r^L)$ . Calling  $g_t := \mathbb{E} \mathcal{W}_2^2(\bar{U}_t^{i,L}, f_t^L)$ , using the Cauchy–Schwarz inequality and the fact that both the particles and the cutoff nonlinear processes have uniformly bounded second moment, we obtain  $\partial_t h_t \leq C[h_t^{1/2} g_t^{1/2} + g_t + L^{1-2/\nu}]$  for almost every  $t \geq 0$ . Using a version of Gronwall’s lemma (see for instance Lemma 4.1.8 of [2]) we deduce that  $h_t \leq C[h_0 + t(1+t)g_t + tL^{1-2/\nu}]$ , and consequently,

$$h_t \leq C[h_0 + t(1+t)N^{-1/3} + tL^{1-2/\nu}],$$

where we have used Corollary 14 to bound  $g_t \leq CN^{-1/3}$  uniformly on  $t$  and  $L$ . From this, we obtain for all  $L \geq 1$

$$\begin{aligned} \mathbb{E} \mathcal{W}_2^2(\bar{V}_t, f_t) &\leq C[\mathbb{E} \mathcal{W}_2^2(\bar{V}_t, \bar{U}_t^L) + \mathbb{E} \mathcal{W}_2^2(\bar{U}_t^L, f_t^L) + \mathcal{W}_2^2(f_t^L, f_t)] \\ &\leq C[h_t + N^{-1/3} + tL^{1-2/\nu}] \\ &\leq C[h_0 + (1+t)^2 N^{-1/3} + tL^{1-2/\nu}], \end{aligned}$$

where we have used Corollary 14 again, together with Lemma 15. Letting  $L \rightarrow \infty$ , the result follows.  $\square$

### 5. Uniform Propagation of Chaos

In this section we give the proof of Theorem 2.

The following is a version of Povzner’s lemma [24], see for instance [6, 15, 23, 33] and the references therein for other versions. Plainly, it will be crucial to establish the propagation of moments for the particle system, needed to take full advantage of the stability result in [25]. A proof can be found in [23] in a slightly different setting; for the readers convenience, we provide a proof of the precise statement below in the Appendix in the case of  $p$  even (which is enough for purposes).

**Lemma 16** (a version of Povzner’s lemma). *Assume that  $I := \int_0^{\pi/2} \theta^2 \beta(\theta) d\theta < \infty$ . Then, for any  $p > 2$  and any  $v, v_* \in \mathbb{R}^3$ , we have*

$$\begin{aligned} &\int_{\mathbb{S}^2} (|v'|^p + |v'_*|^p - |v|^p - |v_*|^p) B(\theta) d\sigma \\ &\leq -A_p (|v|^p + |v_*|^p) + I \tilde{A}_p (|v|^{p-2} |v_*|^2 + |v_*|^{p-2} |v|^2), \end{aligned}$$

where  $\tilde{A}_p > 0$  is some constant that depends only on  $p$ , and

$$A_p := \int_0^{\pi/2} [1 - \sin(\theta/2)^p - \cos(\theta/2)^p] \beta(\theta) d\theta > 0.$$

The propagation of moments for the particle system was already established in [22, Lemma 5.3], where it is assumed that the initial energy is a.s. bounded. In the Maxwellian setting, following the proof of that lemma and performing a careful inspection of how the constants depend on the energy, we are able to write a moments estimate where this dependence is explicit. More importantly, using conditional expectations we are able to deduce propagation of moments for a general initial condition, without any restriction (besides of course finite second moment). This is stated in the next result:

**Corollary 17** (propagation of moments for the particle system). *Assume (5), let  $p \geq 2$  and define the random variable  $\mathcal{E} := \frac{1}{N} \sum_i^N |V_0^i|^2$ . Then, there exists a constant  $C_p > 0$  (nonrandom and not depending on  $\mathbf{V}_0$  nor on its law) such that for all  $t \geq 0$ ,*

$$\mathbb{E} \left( |V_t^1|^p \mid \mathcal{E} \right) \leq \mathbb{E} \left( |V_0^1|^p \mid \mathcal{E} \right) + C_p \mathcal{E}^{p/2}, \quad a.s.$$

As a consequence, we have

$$\mathbb{E}|V_t^1|^p \leq (C_p + 1) \mathbb{E}|V_0^1|^p, \quad \forall t \geq 0.$$

*Proof.* For each  $t \geq 0$ , write  $g_t = \mathbb{E} \left( |V_t^1|^p \mid \mathcal{E} \right)$  which is a.s. equal to  $\mathbb{E} \left( \frac{1}{N} \sum_i^N |V_t^i|^p \mid \mathcal{E} \right)$  by exchangeability of the system  $\mathbf{V}$ . Since the process

$$\begin{aligned} & \frac{1}{N} \sum_i^N |V_t^i|^p - \frac{1}{N} \sum_i^N |V_0^i|^p \\ & - \frac{1}{N} \sum_i^N \int_0^t \int_0^\infty \int_0^{2\pi} \int_{A^i} [|V_s^i + c(V_s^i, V_s^{i(\xi)}, z, \phi)|^p - |V_s^i|^p] \frac{d\xi d\phi dz ds}{2(N-1)\pi} \end{aligned}$$

is a martingale in the filtration  $(\mathcal{F}_t)_{t \geq 0}$  defined in Sect. 2.1 and  $\mathcal{E}$  is  $\mathcal{F}_0$  measurable, taking into account Eq. (8) we get for each  $t \geq 0$

$$\begin{aligned} g_t - g_0 &= \mathbb{E} \left( \frac{1}{N} \sum_i^N \int_0^t \int_0^\infty \int_0^{2\pi} \int_{A^i} [|V_t^i + c(V_t^i, V_t^{i(\xi)}, z, \phi)|^p \right. \\ & \quad \left. - |V_t^i|^p] \frac{d\xi d\phi dz dt}{2(N-1)\pi} \mid \mathcal{E} \right) \\ &= \int_0^t \mathbb{E} \left( \frac{1}{N^2} \sum_i^N \sum_{j \neq i}^N \int_0^{2\pi} \int_0^{\pi/2} [|V_t^i|^p + |V_t^j|^p - |V_t^i|^p \right. \\ & \quad \left. - |V_t^j|^p] \beta(\theta) \frac{d\theta d\phi}{4\pi} \mid \mathcal{E} \right) dt \end{aligned}$$

almost surely. The latter implies that  $t \mapsto g_t$  has an a.s. absolutely continuous version which we work with from now on. Taking the difference  $g_t - g_s$ , dividing by  $t - s$  and letting  $s \rightarrow t$ , by Lemma 16 we a.s. have for some positive constants  $I, A_p$  and  $\tilde{A}_p$  and almost every  $t \geq 0$ :

$$\begin{aligned} \partial_t g_t &\leq -A_p g_t + I \tilde{A}_p \mathbb{E} \left( \frac{1}{N} \sum_i^N |V_t^i|^{p-2} \frac{1}{N-1} \sum_{j \neq i}^N |V_t^j|^2 \mid \mathcal{E} \right) \\ &\leq -A_p g_t + 2\mathcal{E} I \tilde{A}_p g_t^{1-2/p}, \end{aligned}$$

where we have used that  $\sum_{k \neq j} |V_t^k|^2 \leq \sum_k |V_t^k|^2 = N\mathcal{E}$  since the system almost surely preserves energy, together with exchangeability and the conditional Hölder inequality. This differential inequality implies that  $g_t \leq \max(g_0, x^*)$  for all  $t \geq 0$ , where  $x^* =$

$(2\mathcal{E}I\tilde{A}_p/A_p)^{p/2}$  is the unique positive root of the polynomial  $-A_p x + 2\mathcal{E}I\tilde{A}_p x^{1-2/p}$ . This implies

$$g_t \leq g_0 + C_p \mathcal{E}^{p/2},$$

for some constant depending only on  $A_p, \tilde{A}_p$  and  $I$ , which proves the first statement. For the second one, we use conditional expectation to get

$$\begin{aligned} \mathbb{E}|V_t^1|^p &\leq \mathbb{E}\left[\mathbb{E}\left(|V_0^1|^p \mid \mathcal{E}\right) + C_p \mathcal{E}^{p/2}\right] \\ &= \mathbb{E}|V_0^1|^p + C_p \mathbb{E}\left[\left(\frac{1}{N} \sum_{k=1}^N |V_0^k|^2\right)^{p/2}\right], \end{aligned}$$

and then Jensen’s inequality (applied in the empirical mean) gives  $\mathbb{E}|V_t^1|^p \leq \mathbb{E}|V_0^1|^p + C_p \mathbb{E}\frac{1}{N} \sum_{k=1}^N |V_0^k|^p$ . The proof is complete.  $\square$

The fact that the particles have bounded moments allows us to obtain a convergence to equilibrium result for the particle system that extends Proposition 1.7 of [25]. We will state it in terms of the following distance: define  $\mathscr{W}_2$  as the usual 2-Wasserstein distance on  $\mathcal{P}(\mathcal{P}(\mathbb{R}^3))$  induced by  $\mathcal{W}_2$ , that is

$$\mathscr{W}_2(\alpha, \beta) = \inf_{\mathbf{a}, \mathbf{b}} \left(\mathbb{E}\mathcal{W}_2^2(\mathbf{a}, \mathbf{b})\right)^{1/2},$$

where the infimum is taken over all random elements  $\mathbf{a}$  and  $\mathbf{b}$  in  $\mathcal{P}(\mathbb{R}^3)$  such that  $\mathcal{L}(\mathbf{a}) = \alpha$  and  $\mathcal{L}(\mathbf{b}) = \beta$ . For  $F \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$ , denote by  $\hat{F}$  the push-forward of  $F$  by the “empirical measure” map, that is,  $\hat{F} = \mathcal{L}(\bar{\mathbf{X}}) \in \mathcal{P}(\mathcal{P}(\mathbb{R}^3))$  for  $\mathbf{X} \sim F$ . It is also clear that for any  $\mathbf{X} \sim F$  and  $\mu \in \mathcal{P}(\mathbb{R}^3)$ , we have

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{X}}, \mu) = \mathscr{W}_2^2(\hat{F}, \delta_\mu). \tag{29}$$

Define also the *Boltzmann sphere*

$$\mathcal{S}^N = \left\{ \mathbf{x} \in (\mathbb{R}^3)^N : \frac{1}{N} \sum_{i=1}^N x^i = 0, \frac{1}{N} \sum_{i=1}^N |x^i|^2 = 1 \right\},$$

and notice that if  $G_0^N$  is concentrated on  $\mathcal{S}^N$ , then the preservation of momentum and kinetic energy of the collisions imply that  $G_t^N := \mathcal{L}(\mathbf{V}_t)$  is also concentrated on  $\mathcal{S}^N$  for all  $t \geq 0$ . Denote  $\mathcal{U}^N$  the uniform distribution on  $\mathcal{S}^N$ . Theorem 1.6 of [25] states that for any  $\delta > 0$  and  $q > 1$ ,

$$\partial_t^+ \mathscr{W}_2(\hat{G}_t^N, \hat{\mathcal{U}}^N) \leq -c_{\delta,q,N}(G_t^N) \mathscr{W}_2(\hat{G}_t^N, \hat{\mathcal{U}}^N)^{1+1/\delta}, \tag{30}$$

where  $c_{\delta,q,N}(G_t^N) = k_{\delta,q} \mathbb{E}(|V_t^1|^{2q(1+\delta)})^{-1/2q\delta}$  and  $k_{\delta,q} > 0$  is some constant. We are now ready to state and prove:



**Lemma 18.** Assume (5) and that  $G_0^N \in \mathcal{P}_2^{\text{sym}}((\mathbb{R}^3)^N)$  is concentrated on the Boltzmann sphere  $S^N$ . Assume also that  $R_p := \sup_N \mathbb{E}|V_0^1|^p < \infty$  for some  $p \geq 4$ . Then, for all  $0 < \delta < p - 2$  there exists a constant  $C_{p,\delta} > 0$  depending only on  $p, \delta$  and  $R_p$ , such that for all  $N \in \mathbb{N}$  and all  $t \geq 0$

$$\mathcal{W}_2^2(\hat{G}_t^N, \hat{U}^N) \leq \left[ \mathcal{W}_2^2(\hat{G}_0^N, \hat{U}^N)^{-1/\delta} + C_{p,\delta} t \right]^{-\delta}.$$

*Proof.* Take  $q > 1$  and  $\delta < p/2 - 1$  such that  $2q(1 + \delta) = p$ . Using Corollary 17, we deduce that  $c_{\delta,q,N}(G_t^N) \geq k_{\delta,q}(C_p \mathbb{E}|V_0^1|^p)^{-1/2q\delta} \geq k_{\delta,q}(C_p R_p)^{-1/2q\delta} =: C_{p,\delta}$ . From (30), the result follows using Gronwall’s lemma, squaring, and redefining  $\delta$  as  $2\delta$ .  $\square$

We conclude with the proof of Theorem 2. We follow a standardization argument found in [10], which allows one to reduce the proof to the case where the initial distribution  $G_0^N$  is concentrated on the Boltzmann sphere  $S^N$ .

Given any  $F \in \mathcal{P}^{\text{sym}}((\mathbb{R}^3)^N)$  and a random vector  $\mathbf{X} \sim F$ , set  $[F] = \mathcal{L}([\mathbf{X}])$ , where  $[\mathbf{X}] = \mathbf{Y} = (Y^1, \dots, Y^N)$  is defined as

$$Y^i = \frac{X^i - M}{S}, \quad \text{with } M = \frac{1}{N} \sum_j X^j, \quad S^2 = \frac{1}{N} \sum_j |X^j - M|^2 \tag{31}$$

on the event  $\{S > 0\}$ , and  $Y^i = Z^i$  on  $\{S = 0\}$ , where  $\mathbf{Z}$  is some (arbitrary but fixed) exchangeable random vector on  $S^N$ , independent of  $\mathbf{X}$  (although we will mainly use this standardization when  $S > 0$ ). This, of course, ensures that  $[F]$  is symmetric and concentrated on  $S^N$ . Calling  $Q^2 = \frac{1}{N} \sum_i |X^i|^2$  we have  $S^2 = Q^2 - |M|^2$ , and then

$$\begin{aligned} \frac{1}{N} \sum_i |X^i - Y^i|^2 &= \mathbf{1}_{\{S>0\}} \frac{1}{N} \sum_i \left| \frac{(S-1)X^i + M}{S} \right|^2 + \mathbf{1}_{\{S=0\}} \frac{1}{N} \sum_i |M - Z^i|^2 \\ &= \mathbf{1}_{\{S>0\}} \frac{(S-1)^2 Q^2 + (2S-1)|M|^2}{S^2} + \mathbf{1}_{\{S=0\}} [1 + |M|^2] \\ &= \mathbf{1}_{\{S>0\}} [(S-1)^2 + |M|^2] + \mathbf{1}_{\{S=0\}} [1 + |M|^2] \\ &= (S-1)^2 + |M|^2 \leq \mathcal{W}_2^2(\bar{\mathbf{X}}, \mu), \end{aligned} \tag{32}$$

where  $\mu \in \mathcal{P}_2(\mathbb{R}^3)$  is any distribution with  $M_\mu := \int v \mu(dv) = 0$  and  $S_\mu^2 := \int |v - M_\mu|^2 \mu(dv) = 1$  (in general,  $(S_\mu - S_\nu)^2 + |M_\mu - M_\nu|^2 \leq \mathcal{W}_2^2(\mu, \nu)$ , since for  $X \sim \mu$  and  $Y \sim \nu$  one has  $\mathbb{E}|X - Y|^2 = \mathbb{E}[|(X - M_\mu) - (Y - M_\nu)|^2] + |M_\mu - M_\nu|^2 \geq S_\mu^2 + S_\nu^2 - 2S_\mu S_\nu + |M_\mu - M_\nu|^2$ ). Since  $\mathcal{W}_2^2(F, [F]) \leq \mathbb{E} \frac{1}{N} \sum_i |X^i - Y^i|^2$ , this gives for any such  $\mu$ :

$$\mathcal{W}_2^2(F, [F]) \leq \mathbb{E} \mathcal{W}_2^2(\bar{\mathbf{X}}, \mu) = \mathcal{W}_2^2(\hat{F}, \delta_\mu). \tag{33}$$

*Remark 19.* If  $\gamma$  denotes the Gaussian density with mean 0 and variance 1, that is,  $\gamma(v) = (2\pi\sigma^2)^{-3/2} e^{-|v|^2/(2\sigma^2)}$  for  $\sigma^2 = 1/3$ , then the measure  $[\gamma^{\otimes N}]$  corresponds to  $\mathcal{U}^N$ . Thus, from (33) applied to  $F = \gamma^{\otimes N}$  and  $\mu = \gamma$ , we obtain  $\mathcal{W}_2^2(\gamma^{\otimes N}, \mathcal{U}^N) \leq \mathcal{W}_2^2(\widehat{\gamma^{\otimes N}}, \delta_\gamma)$ . Since  $\mathcal{W}_2^2(\widehat{\gamma^{\otimes N}}, \hat{U}^N) \leq \mathcal{W}_2^2(\gamma^{\otimes N}, \mathcal{U}^N)$  and using (28) on  $\mathcal{W}_2^2(\widehat{\gamma^{\otimes N}}, \delta_\gamma) = \varepsilon_N(\gamma)$ , we deduce the following chaos rate for  $\mathcal{U}^N$  (already established in [10, Lemma 25-(i)]; see [18, Theorem 4.4] for related  $\mathcal{W}_1$  estimates):

$$\mathcal{W}_2^2(\hat{U}^N, \delta_\gamma) \leq 2\mathcal{W}_2^2(\hat{U}^N, \widehat{\gamma^{\otimes N}}) + 2\mathcal{W}_2^2(\widehat{\gamma^{\otimes N}}, \delta_\gamma) \leq 4\mathcal{W}_2^2(\widehat{\gamma^{\otimes N}}, \delta_\gamma) \leq CN^{-1/2}. \tag{34}$$

*Proof of Theorem 2.* Let us first prove the result in the case where  $G_0^N$  is concentrated on the Boltzmann sphere  $\mathcal{S}^N$ . Noting that in this case  $\mathcal{W}_2^2(\hat{G}_0^N, \hat{U}^N) \leq 4$ , Lemma 18 gives

$$\mathcal{W}_2^2(\hat{G}_t^N, \hat{U}^N) \leq C_{p,\delta}(1+t)^{-\delta}, \tag{35}$$

for all  $\delta < p - 2$ , where  $C_{p,\delta}$  depends only on  $p, \delta$  and  $R_p := \sup_N \mathbb{E}|V_0^1|^p$ . With this, from (29) we have:

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) = \mathcal{W}_2^2(\hat{G}_t^N, \delta_{f_t}) \leq C \left[ \mathcal{W}_2^2(\hat{G}_t^N, \hat{U}^N) + \mathcal{W}_2^2(\hat{U}^N, \delta_\gamma) + \mathcal{W}_2^2(\delta_\gamma, \delta_{f_t}) \right].$$

The first and second terms are controlled using (35) and (34), respectively. The third term is equal to  $\mathcal{W}_2^2(\gamma, f_t)$ , which, by Theorem 5.8 of [14], converges exponentially fast to 0 under our assumption that  $f_0$  has finite  $p_0$ -moment for some  $p_0 > 4$  (condition (6)). Thus, we can just bound that term by  $C_{p,\delta}(1+t)^{-\delta}$ . All this yields

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C_{p,\delta}(1+t)^{-\delta} + CN^{-1/3}.$$

For  $\eta \in (0, 1/3)$  to be chosen, set  $\bar{t} = N^{\eta/\delta} - 1$ , so the last inequality implies that  $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq C_{p,\delta}N^{-\eta}$  for all  $t > \bar{t}$ , whereas Theorem 1 gives  $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq Ch_0 + CN^{-1/3+2\eta/\delta}$  for all  $t \leq \bar{t}$ , where  $h_0 = \mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$ . Setting  $\eta = [3(1 + 2/\delta)]^{-1} < (p - 2)/3p$  gives  $\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq Ch_0 + C_{p,\delta}N^{-\eta}$  for all  $t \geq 0$ . Since  $\delta$  can be chosen arbitrarily close to  $p - 2$ , the result follows in this case.

Now we prove the general case by reducing it to the previous one. Consider the process  $\mathbf{W}_t := [\mathbf{V}_t]$ , with the notation of (31). Set  $M_t = \frac{1}{N} \sum_j V_t^j = M_0$ ,  $S_t^2 = \frac{1}{N} \sum_j |V_t^j - M_0|^2 = S_0^2$ . It holds that on the event  $\{S_0 > 0\}$ ,  $\mathbf{W}$  solves (13) with the same Poisson measure associated to  $\mathbf{V}$ , but starting with initial condition  $\mathbf{W}_0 = [\mathbf{V}_0]$ . Specifically: given  $v, v_*, m \in \mathbb{R}^3$  and  $s > 0$ , the homogeneity of  $\hat{v}$  and  $\hat{j}$  and the definitions of (7) imply that for  $\tilde{v} = (v - m)/s$  and  $\tilde{v}_* = (v_* - m)/s$  we have  $c(\tilde{v}, \tilde{v}_*, z, \phi) = c(v, v_*, z, \phi)/s$ , and then

$$\tilde{v}' = \tilde{v} + c(\tilde{v}, \tilde{v}_*, z, \phi) = \frac{v + c(v, v_*, z, \phi) - m}{s} = \frac{v' - m}{s},$$

and the same for  $\tilde{v}'_*$ . This means that the standardization procedure is preserved by the collisions. Since the function  $c$  is the one involved in the definition of the particle system, this shows that  $\mathbf{W}$  solves (13), as desired.

Define the event  $D = \{S_0^2 \geq 1/4\}$ . We have

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq 2\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \bar{\mathbf{W}}_t) + 2\mathbb{E}\mathbf{1}_D \mathcal{W}_2^2(\bar{\mathbf{W}}_t, f_t) + \mathbb{E}\mathbf{1}_{D^c} \mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t). \tag{36}$$

From (32) we have  $\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \bar{\mathbf{W}}_t) \leq \frac{1}{N} \sum_i |V_t^i - W_t^i|^2 = (S_t - 1)^2 + |M_t|^2 = (S_0 - 1)^2 + |M_0|^2 \leq \mathcal{W}_2^2(\bar{\mathbf{V}}_0, f_0)$ , since  $\int v f_0(dv) = 0$  and  $\int |v|^2 f_0(dv) = 1$ . Thus,  $\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \bar{\mathbf{W}}_t) \leq \mathcal{W}_2^2(\bar{\mathbf{V}}_0, f_0) \leq 2\mathcal{W}_2^2(\bar{\mathbf{V}}_0, \bar{\mathbf{U}}_0) + 2\mathcal{W}_2^2(\bar{\mathbf{U}}_0, f_0)$ , and then for the first term of (36) we have

$$\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_t, \bar{\mathbf{W}}_t) \leq 2\mathbb{E} \frac{1}{N} \sum_i |V_0^i - U_0^i|^2 + 2\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{U}}_0, f_0) \leq 2\mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + CN^{-1/2}, \tag{37}$$

where we have used (28).

On  $D$ , we have  $|W_0^i| = |V_0^i + M_0|/S_0 \leq 2(|V_0^i| + |M_0|)$ , and thus  $\mathbb{E}(\mathbf{1}_D |W_0^i|^p) \leq C_p \mathbb{E}|V_0^i|^p$ . Since  $\mathbf{W}$  is a particle system taking values on  $\mathcal{S}^N$ , we can apply the previous case, obtaining  $\mathbb{E} \mathbf{1}_D \mathcal{W}_2^2(\bar{\mathbf{W}}_t, f_t) \leq C \mathcal{W}_2^2([G_0^N], f_0^{\otimes N}) + C_{p,\epsilon} N^{-(p-2)/3p+\epsilon}$ . Using (33) we have  $\mathcal{W}_2^2([G_0^N], f_0^{\otimes N}) \leq 2\mathcal{W}_2^2([G_0^N], G_0^N) + 2\mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) \leq 2\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_0, f_0) + 2\mathcal{W}_2^2(G_0^N, f_0^{\otimes N})$ , and from there, the same argument used to estimate the first term of (36) yields for the second term:

$$\mathbb{E} \mathbf{1}_D \mathcal{W}_2^2(\bar{\mathbf{W}}_t, f_t) \leq C \mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + C_{p,\epsilon} N^{-(p-2)/3p+\epsilon}. \tag{38}$$

For the third term, using the preservation of momentum and energy, we have

$$\begin{aligned} \mathbb{E} \mathbf{1}_{D^c} \mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) &\leq \mathbb{E} \left( \mathbf{1}_{\{S_0^2 < 1/4\}} 2 \left[ \frac{1}{N} \sum_i |V_t^i|^2 + \int |v|^2 f_t(dv) \right] \right) \\ &= \mathbb{E} \left( \mathbf{1}_{\{S_0^2 < 1/4\}} 2[S_0^2 + |M_0|^2 + 1] \right) \\ &\leq \frac{10}{4} \mathbb{P}(S_0^2 < 1/4) + 2\mathbb{E}|M_0|^2, \end{aligned}$$

and since  $\mathbb{P}(S_0^2 < 1/4) \leq \mathbb{P}(|S_0 - 1| > 1/2) \leq 4\mathbb{E}|S_0 - 1|^2$ , we have shown that  $\mathbb{E} \mathbf{1}_{D^c} \mathcal{W}_2^2(\bar{\mathbf{V}}_t, f_t) \leq 10\mathbb{E}(|S_0 - 1|^2 + |M_0|^2)$ , which is again controlled by  $10\mathbb{E}\mathcal{W}_2^2(\bar{\mathbf{V}}_0, f_0) \leq C \mathcal{W}_2^2(G_0^N, f_0^{\otimes N}) + CN^{-1/2}$  as above.

Finally, putting the previous estimate, (37) and (38) into (36), yields the desired result.  $\square$

### Appendix

*Proof of Lemma 8.* Part (i) follows from the last step in the proof of part (ii) below.

Since  $L < \infty$ , in order to prove (ii) it is enough to construct a weak solution  $\mathbf{V}$  to (13) and then build  $\mathbf{U}^L$  driven by the same Poisson process in a jump-by-jump manner.

To obtain a weak solution of (13), we will use a cutoff procedure: for a given cutoff level  $K \in [1, \infty)$ , define  $\mathbf{V}^K = (V^{1,K}, \dots, V^{N,K})$  as the solution to

$$d\mathbf{V}_t^K = \int_0^\infty \int_0^{2\pi} \int_{[0,N)^2} \sum_{i \neq j} \mathbf{1}_{\{i(\xi)=i, i(\zeta)=j\}} \mathbf{c}_{K,ij}(\mathbf{V}_{t-}^K, z, \phi) \mathcal{N}(dt, dz, d\phi, d\xi, d\zeta) \tag{39}$$

where  $\mathbf{c}_{K,ij}$  is defined as in (14) but using  $c_K$  in place of  $c$ . Again, since  $K < \infty$ , the system  $\mathbf{V}^K$  can be constructed pathwise. Thus, given a sequence of finite cutoff levels  $K \rightarrow \infty$ , one can prove in a similar way as in Proposition 1.2–(ii) of [13] that the laws of  $\mathbf{V}^K$  are tight (the second moment estimates are indeed trivial here because of their exact preservation stated in (i)). By martingale methods and classic probability space enlargement arguments, one then gets that the accumulation points are weak solutions of (13). In order to prove uniqueness in law of weak solutions of (13) it is enough to show that any weak solution can be approximated, in a pathwise way as  $K \rightarrow \infty$ , by (strong) solutions  $\hat{\mathbf{V}}^K$  to (39) driven by some Poisson measures defined in the same probability space. More specifically: given  $\hat{\mathbf{V}}^\infty$  a weak solution to (13) driven by the Poisson point measure  $\hat{\mathcal{N}}(dt, dz, d\phi, d\xi, d\zeta)$ , define  $\hat{\mathbf{V}}^K$  uniquely by

$$d\hat{\mathbf{V}}_t^K = \int_0^\infty \int_0^{2\pi} \int_{[0,N)^2} \sum_{i \neq j} \mathbf{1}_{\{i(\xi)=i, i(\zeta)=j\}} \\ \times \mathbf{c}_{K,ij}(\hat{\mathbf{V}}_{t-}^K, z, \phi_{t-}^K) \hat{\mathcal{N}}(dt, dz, d\phi, d\xi, d\zeta)$$

with  $\hat{\mathbf{V}}_0^K = \hat{\mathbf{V}}_0$  and where  $\phi_{t-}^K = \phi_{t-}^K(\phi, \xi, \zeta) = \varphi(\hat{\mathbf{V}}_{t-}^{i(\xi),\infty} - \hat{\mathbf{V}}_{t-}^{i(\zeta),\infty}, \hat{\mathbf{V}}_{t-}^{i(\xi),K} - \hat{\mathbf{V}}_{t-}^{i(\zeta),K}, \phi)$ . Notice that  $\hat{\mathbf{V}}^K$  has the same law as  $\mathbf{V}^K$ , by pathwise uniqueness for (13) and the fact that the point measure  $\hat{\mathcal{N}}^K$  defined on test functions by  $f \mapsto \int f(t, z, \phi_{t-}^K, d\xi, d\zeta) \hat{\mathcal{N}}(dt, dz, d\phi, d\xi, d\zeta)$  has the same law as  $\mathcal{N}(dt, dz, d\phi, d\xi, d\zeta)$  (as can be checked using Itô calculus and Campbell’s formula). Since the coupling between circles provided by Lemma 6 is optimal, integrated versions of the bounds provided in Lemma 5.1 of [13] (for some other coupling of angles  $\phi$ ) also hold in the present context. Using Itô calculus and Gronwall’s lemma we are then able to prove, in a similar way that, for each  $T > 0$  and some function  $R_T(K)$  going to 0 as  $K \rightarrow \infty$  one has  $\sup_{t \in [0,T]} \mathbb{E}(|\hat{\mathbf{V}}_t^K - \hat{\mathbf{V}}_t^\infty|^2) \leq R_T(K)$ , from where we get the desired convergence:

$$\mathbb{E} \left( \sup_{t \in [0,T]} |\hat{\mathbf{V}}_t^K - \hat{\mathbf{V}}_t^\infty|^2 \right) \leq R_T(K)$$

Part (ii) follows.

Now we prove part (iii). Define  $\tilde{\mathcal{N}}^{i,L}(dt, dz, d\phi, dv)$  as the point measure on  $[0, \infty) \times [0, \infty) \times [0, 2\pi) \times \mathbb{R}^3$  with atoms  $(t, z, \phi_{t-}^i, \Pi_{t-}^{i,L}(\mathbf{U}_{t-}^L, \xi))$  for every atom  $(t, z, \phi, \xi)$  of  $\mathcal{N}^i$ . Since: (1) the dependence on  $(\mathbf{V}, \mathbf{U}^L)$  is predictable, (2) the  $\xi$ -law of  $\Pi_{t-}^{i,L}(\mathbf{x}, \xi)$  is  $f_t^L$  for any  $\mathbf{x} \in (\mathbb{R}^3)^N$ , and (3) the  $\phi$ -law of  $\varphi(v, u, \phi)$  is the uniform distribution on  $[0, 2\pi)$  for any  $v, u \in \mathbb{R}^3$ , one can compute the Laplace functional of  $\tilde{\mathcal{N}}^{i,L}$  and deduce that it is a Poisson point measure with intensity  $dt dz d\phi f_t^L(dv)/2\pi$ . From (17), it is then clear that  $U^{i,L}$  satisfies (11), with  $\mathcal{M}^L$  replaced by  $\tilde{\mathcal{N}}^{i,L}$ , which shows that  $U^{i,L}$  is a cutoff nonlinear process.

Part (iv) is obvious.  $\square$

*Proof of Lemma 16, case of  $p$  even.* We will use the parametrization of  $\mathbb{S}^2$  given by (7), that is,  $v' = v + a(v, v_*, \theta, \phi)$  and  $v'_* = v_* - a(v, v_*, \theta, \phi)$ . For notational simplicity, call  $c := \frac{1-\cos\theta}{2}$ ,  $s = \frac{\sin\theta}{2}$ ,  $\hat{i} = \hat{i}(v - v_*)$ ,  $\hat{j} = \hat{j}(v - v_*)$  and  $\Gamma = \Gamma(v - v_*, \phi)$  (recall that  $\Gamma = (\cos\phi)\hat{i} + (\sin\phi)\hat{j}$ ). Noting that  $|\Gamma|^2 = |v - v_*|^2$  and that  $(v - v_*) \cdot \Gamma = 0$ , we have

$$|v'|^2 = |v + c(v_* - v) + s\Gamma|^2 \\ = [1 + c^2 + s^2 - 2c]|v|^2 + [c^2 + s^2]|v_*|^2 + 2[c - c^2 - s^2]v \cdot v_* + 2sv \cdot \Gamma \\ = (1 - c)|v|^2 + c|v_*|^2 + s(v + v_*) \cdot \Gamma,$$

where in the last step we used the identity  $c^2 + s^2 = c$  and the fact that  $v \cdot \Gamma = v_* \cdot \Gamma$ . Calling  $w = v + v_*$ , we thus obtain

$$|v'|^2 = x + sw \cdot \Gamma, \quad \text{for } x = (1 - c)|v|^2 + c|v_*|^2, \text{ and} \\ |v'_*|^2 = y - sw \cdot \Gamma, \quad \text{for } y = (1 - c)|v_*|^2 + c|v|^2,$$

where the second identity is deduced similarly as the first one. Take  $p = 2k$  for integer  $k \geq 2$ . Thus,

$$\int_0^{2\pi} |v'|^p \frac{d\phi}{2\pi} = x^k + \sum_{i=1}^{\lfloor k/2 \rfloor} \binom{k}{2i} x^{k-2i} s^{2i} \int_0^{2\pi} (w \cdot \Gamma)^{2i} \frac{d\phi}{2\pi}, \tag{40}$$

where we have used the fact that  $\int_0^{2\pi} (w \cdot \Gamma)^{2i-1} = 0$ , since it is computed as the sum of integrals of terms of the form  $(\cos \phi)^a (\sin \phi)^b$  where  $a$  or  $b$  is odd. The key of the proof is to show that, after integration in  $\phi$ , the term  $(w \cdot \Gamma)^{2i}$  is of order  $|v|^{2i} |v_*|^{2i}$  and not  $|v|^{4i} |v_*|^{4i}$ , as one would obtain using loose bounds. Specifically: using the same argument to neglect the odd terms of the sum, we have

$$\begin{aligned} \int_0^{2\pi} (w \cdot \Gamma)^{2i} \frac{d\phi}{2\pi} &= \int_0^{2\pi} [(\cos \phi)w \cdot \hat{i} + (\sin \phi)w \cdot \hat{j}]^{2i} \frac{d\phi}{2\pi} \\ &= \sum_{j=0}^i \binom{2i}{2j} (w \cdot \hat{i})^{2i-2j} (w \cdot \hat{j})^{2j} \int_0^{2\pi} (\cos \phi)^{2i-2j} (\sin \phi)^{2j} \frac{d\phi}{2\pi}. \end{aligned}$$

Denoting by  $n!!$  the product of the positive integers smaller than or equal to  $n$  which have the same parity as  $n$ , one can check the identity  $\int_0^{2\pi} (\cos \phi)^n (\sin \phi)^m \frac{d\phi}{2\pi} = \frac{(n-1)!!(m-1)!!}{(n+m)!!}$  for  $n$  and  $m$  even (integrate by parts in both possible ways, use  $\cos^2 \phi + \sin^2 \phi = 1$  to get two recurrence relations and deduce the identity by double induction in  $(n, m)$ ). Since  $(2n-1)!! = \frac{(2n)!}{2^n n!}$  and  $(2n)!! = 2^n n!$ , it can be easily seen that  $\binom{2i}{2j} \frac{(2i-2j-1)!!(2j-1)!!}{(2i)!!} = 2^{-2i} \binom{2i}{i} \binom{i}{j}$ , which yields

$$\begin{aligned} \int_0^{2\pi} (w \cdot \Gamma)^{2i} \frac{d\phi}{2\pi} &= 2^{-2i} \binom{2i}{i} \sum_{j=0}^i \binom{i}{j} (w \cdot \hat{i})^{2i-2j} (w \cdot \hat{j})^{2j} \\ &= 2^{-2i} \binom{2i}{i} \left[ (w \cdot \hat{i})^2 + (w \cdot \hat{j})^2 \right]^i. \end{aligned}$$

Using that  $(\frac{v-v_*}{|v-v_*|}, \frac{\hat{i}}{|v-v_*|}, \frac{\hat{j}}{|v-v_*|})$  is an orthonormal basis, we have

$$\begin{aligned} (w \cdot \hat{i})^2 + (w \cdot \hat{j})^2 &= |v - v_*|^2 |w|^2 - (w \cdot (v - v_*))^2 \\ &= (|v|^2 + |v_*|^2 - 2v \cdot v_*) (|v|^2 + |v_*|^2 + 2v \cdot v_*) - (|v|^2 - |v_*|^2)^2 \\ &= 4|v|^2 |v_*|^2 - 4(v \cdot v_*)^2, \end{aligned}$$

where the cancelation of  $|v|^4$  and  $|v_*|^4$  in the last step is the crucial point of the proof. We deduce that

$$\int_0^{2\pi} (w \cdot \Gamma)^{2i} \frac{d\phi}{2\pi} = \binom{2i}{i} \left[ |v|^2 |v_*|^2 - (v \cdot v_*)^2 \right]^i \leq 2^i \binom{2i}{i} |v|^{2i} |v_*|^{2i}$$

Denote  $\tilde{A}_p > 0$  some constant that depends only on  $p$  and that can change from line to line. With the last inequality and using the bound  $x^{k-2i} \leq \tilde{A}_p (|v|^{2k-4i} + |v_*|^{2k-4i})$ , from (40) we obtain

$$\begin{aligned} \int_0^{2\pi} |v'|^p \frac{d\phi}{2\pi} &\leq c^k |v|^{2k} + (1-c)^k |v_*|^{2k} + \sum_{i=1}^{k-1} \binom{k}{i} (1-c)^{k-i} |v|^{2k-2i} c^i |v_*|^{2i} \\ &\quad + \tilde{A}_p \sum_{i=1}^{\lfloor k/2 \rfloor} s^{2i} (|v|^{2k-2i} |v_*|^{2i} + |v_*|^{2k-2i} |v|^{2i}) \\ &\leq c^k |v|^{2k} + (1-c)^k |v_*|^{2k} + \theta^2 \tilde{A}_p (|v|^{2k-2} |v_*|^2 + |v_*|^{2k-2} |v|^2), \end{aligned}$$

where in the last step we have used Young’s inequality and the fact that  $s^{2i}$  and  $c^i$  are of order  $\theta^2$  for  $i \geq 1$ . With the same argument, the last inequality is valid replacing  $v'$  for  $v'_*$  and exchanging the roles of  $v$  and  $v_*$ . Since  $c = \sin(\theta/2)^2$  and  $1-c = \cos(\theta/2)^2$ , integrating  $|v'|^p + |v'_*|^p - |v|^p - |v_*|^p$  against  $\beta(\theta)d\theta \frac{d\phi}{2\pi}$  yields the result.  $\square$

*Acknowledgements.* R.C.’s research was partially supported by Mecesup UCH0607 Doctoral Fellowship, BASAL-Conicyt Center for Mathematical Modeling (CMM), ICM Millennium Nucleus NC120062 and Fondecyt Postdoctoral Project 3160250. J.F. acknowledges partial support from Fondecyt Project 1150570, ICM Millennium Nucleus NC120062 and BASAL-Conicyt CMM. Both authors thank Nicolas Fournier for exchanges of mutual interest. We also thank anonymous referees for pointing out to us the reference [3], for suggesting the use of the distance  $\mathscr{H}_2$  to simplify the proof of Theorem 2 and for other suggestions that allowed us to improve the presentation of the paper.

**References**

1. Alexandre, R.: A review of Boltzmann equation with singular kernels. *Kinet. Relat. Models* **2**(4), 551–646 (2009)
2. Ambrosio, L., Gigli, N., Savaré, G.: Gradient flows in metric spaces and in the space of probability measures, 2nd edn. *Lectures in Mathematics ETH Zürich*. Birkhäuser Verlag, Basel (2008)
3. Carrapatoso, K.: Quantitative and qualitative Kac’s chaos on the Boltzmann’s sphere. *Ann. Inst. Henri Poincaré Probab. Stat.* **51**(3), 993–1039 (2015)
4. Cercignani, C.: *The Boltzmann Equation and Its Applications*, vol. 67 of Applied Mathematical Sciences. Springer, New York (1988)
5. Cortez, R., Fontbona, J.: Quantitative propagation of chaos for generalized Kac particle systems. *Ann. Appl. Probab.* **26**(2), 892–916 (2016)
6. 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)
7. Fontbona, J., Guérin, H., Méléard, S.: Measurability of optimal transportation and convergence rate for Landau type interacting particle systems. *Probab. Theory Relat. Fields.* **143**(3–4), 329–351 (2009)
8. Fontbona, J., Guérin, H., Méléard, S.: Measurability of optimal transportation and strong coupling of martingale measures. *Electron. Commun. Probab.* **15**, 124–133 (2010)
9. Fournier, N.: Finiteness of entropy for the homogeneous Boltzmann equation with measure initial condition. *Ann. Appl. Probab.* **25**(2), 860–897 (2015)
10. Fournier, N., Guillin, A.: From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules. Preprint, [arXiv:1510.01123](https://arxiv.org/abs/1510.01123) (2015)
11. Fournier, N., Guillin, A.: On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Relat. Fields* **162**(3–4), 707–738 (2015)
12. Fournier, N., Méléard, S.: A stochastic particle numerical method for 3D Boltzmann equations without cutoff (electronic). *Math. Comput.* **71**(238), 583–604 (2002)
13. 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)
14. Gabetta, G., Toscani, G., Wennberg, B.: Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation. *J. Stat. Phys.* **81**(5–6), 901–934 (1995)
15. Gamba, I. M., Panferov, V., Villani, C.: On the Boltzmann equation for diffusively excited granular media. *Commun. Math. Phys.* **246**(3), 503–541 (2004)
16. Graham, C., Méléard, S.: Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *Ann. Probab.* **25**(1), 115–132 (1997)
17. Grünbaum F, A.: Propagation of chaos for the Boltzmann equation. *Arch. Ration. Mech. Anal.* **42**, 323–345 (1971)

18. Hauray, M., Mischler, S.: On Kac's chaos and related problems. *J. Funct. Anal.* **266**(10), 6055–6157 (2014)
19. Kac, M.: Foundations of kinetic theory. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955*, vol. III (Berkeley and Los Angeles), pp. 171–197 University of California Press (1956)
20. McKean H., P. Jr.: An exponential formula for solving Boltzmann's equation for a Maxwellian gas. *J. Comb. Theory* **2**, 358–382 (1967)
21. Méléard, S.: Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In: *Probabilistic Models for Nonlinear Partial Differential Equations* (Montecatini Terme, 1995), vol. 1627 of *Lecture Notes in Mathematics*, pp. 42–95. Springer, Berlin (1996)
22. Mischler, S., Mouhot, C.: Kac's program in kinetic theory. *Invent. Math.* **193**(1), 1–147 (2013)
23. Mischler, S., Wennberg, B.: On the spatially homogeneous Boltzmann equation. *Ann. Inst. Henri Poincaré Anal. Non Linéaire* **16**(4), 467–501 (1999)
24. Povzner A., J.: On the Boltzmann equation in the kinetic theory of gases. *Mat. Sb. (N.S.)*. **58**(100), 65–86 (1962)
25. Rousset, M.: A  $N$ -uniform quantitative Tanaka's theorem for the conservative Kac's  $N$ -particle system with Maxwell molecules. Preprint, [arXiv:1407.1965](https://arxiv.org/abs/1407.1965) (2014)
26. Sznitman, A.-S.: Équations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebiete*. **66**(4), 559–592 (1984)
27. Sznitman, A.-S.: Topics in propagation of chaos. In: *École d'Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of *Lecture Notes in Mathematics*, pp. 165–251. Springer, Berlin (1991)
28. Tanaka, H.: On the uniqueness of Markov process associated with the Boltzmann equation of Maxwellian molecules. In: *Proceedings of the International Symposium on Stochastic Differential Equations* (Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 1976), pp. 409–425. Wiley, New York (1978)
29. Tanaka, H.: Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete*. **46**(1), 67–105 (1978/1979)
30. Toscani, G., Villani, C.: Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas. *J. Stat. Phys.* **94**(3–4), 619–637 (1999)
31. Villani, C.: A review of mathematical topics in collisional kinetic theory. In: *Handbook of Mathematical Fluid Dynamics*, vol. I, pp. 71–305. North-Holland, Amsterdam (2002)
32. Villani, C.: *Optimal Transport, Old and New*, vol. 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Berlin (2009)
33. Wennberg, B.: Entropy dissipation and moment production for the Boltzmann equation. *J. Stat. Phys.* **86**(5–6), 1053–1066 (1997)

Communicated by C. Mouhot