SIAM J. MATH. ANAL.
Vol. 56, No. 4, pp. 5409–5444

DERIVATION OF A BOLTZMANN EQUATION WITH HIGHER-ORDER COLLISIONS FROM A GENERALIZED KAC MODEL*

ESTEBAN CÁRDENAS†, NATAŠA PAVLOVIĆ†, AND WILLIAM WARNER†

Abstract. In this work, we generalize Kac's original many-particle binary stochastic model to derive a space homogeneous Boltzmann equation that includes a linear combination of higher-order collisional terms. First, we prove an abstract theorem about convergence from a finite hierarchy to an infinite hierarchy of coupled equations. We apply this convergence theorem on hierarchies for marginals corresponding to the generalized Kac model mentioned above. As a corollary, we prove propagation of chaos for the marginals associated to the generalized Kac model. In particular, the first marginal converges towards the solution of a Boltzmann equation including interactions up to a finite order and whose collision kernel is of Maxwell type with cut-off.

Key words. Boltzmann equation, Kac model, higher-order collisions

MSC codes. 82C40, 82C22

DOI. 10.1137/23M1606150

1. Introduction. The aim of this paper is to derive a Boltzmann equation for Maxwell molecules that incorporates higher-order collisions; we achieve that by generalizing Kac's original stochastic binary model [28] via allowing multiparticle interactions. With this purpose in mind, let us consider a space homogeneous gas of indistinguishable particles, moving in d-dimensional Euclidean space. The system is to be described by the probability density $f = f(t, v)$ of finding a single particle with velocity $v \in \mathbb{R}^d$ at time $t \geq 0$. The resulting Boltzmann-type equation will be of the form

$$\partial_t f = \beta_1 Q_1(f) + \beta_2 Q_2(f, f) + \cdots + \beta_M Q_M(f, \dots, f), \tag{1.1}$$

where $(\beta_K)_{K=1}^M$ is a normalized set of coefficients: $\sum \beta_K = 1$. Here, $M \in \mathbb{N}$ is the highest-order collision that will be relevant in our system, and $(Q_K)_{K=1}^M$ are the Kth collisional operators, modeling the interactions between K particles.

Since the Boltzmann equation was introduced by Boltzmann [5] and Maxwell [30], it has been the target of many mathematical studies. In particular, the problem of rigorously deriving a Boltzmann equation with *binary* interactions (of Maxwell type) was first addressed by Kac in his foundational work [28]. By setting up an appropriate N-particle stochastic process, Kac was able to show that an equation of the form (1.1)—with the right-hand side containing only the Q_2 term—emerges from the many-particle dynamics in the $N \to \infty$ limit. The framework introduced in [28] is now known as the *Kac model,* and there is an active field of research around it; its simplicity is a fertile playground for studying subtle questions that are otherwise very difficult to approach

*Received by the editors October 2, 2023; accepted for publication (in revised form) May 6, 2024; published electronically July 30, 2024.
https://doi.org/10.1137/23M1606150
Funding: The first author was supported by the Provost's Graduate Excellence Fellowship at The University of Texas at Austin and by NSF grant DMS-2009549. The second author was supported by the NSF under grants DMS-1840314, DMS-2009549, and DMS-2052789. The third author was supported by NSF grants DMS-1840314, DMS-2009549, and DMS-2052789.
†Department of Mathematics, University of Texas at Austin, Austin, TX 78712 USA (eacardenas@utexas.edu, natasa@math.utexas.edu, billy.warner@utexas.edu).

in more complex models arising from kinetic theory. *Propagation of chaos, entropy production, relaxation towards equilibrium*, and *well-posedness* are among the most studied questions for the Kac model and its generalizations. For a partial survey of articles, see, e.g., [28, 33, 34, 7, 16, 17, 18, 13, 14, 26, 39, 20, 6, 19, 27, 10, 11, 31, 9, 12, 25] and references therein.

Derivation of the Boltzmann equation in the deterministic, space-inhomogeneous setting with hard spheres has been a major breakthrough in kinetic theory. The first proof in this direction was given by Lanford [29]. More recently, this derivation program has been revisited in a modern perspective by Gallagher, Saint-Raymond, and Texier [23]. On the other hand, derivations of Boltzmann-type equations that include *higher-order* collisions between the particles has just recently started to receive more attention. In [1], I. Ampatzoglou and the second author of this paper derived the nonhomogeneous Boltzmann equation for hard spheres, with the relevant interactions being ternary. In [2], the same authors were able to simultaneously include both binary and ternary interactions in their analysis. The problem of including arbitrarily higher-order interactions remains open. We would also like to point out the recent work [3], which implies that an equation of the type (1.1) including a linear combination of collision operators can give better properties of solutions compared to the binary Boltzmann equation. Specifically, in [3], Ampatzoglou, Gamba, Tasković, and the second author of this paper have shown that the simultaneous existence of binary and ternary collisions in a homogeneous Boltzmann-type equation yields better generation in time properties of moments and time decay, compared to when only binary or ternary collisions are considered. This gives additional motivation to study both derivation as well as analysis of Boltzmann equations with higher-order collisions such as (1.1), which is what we do in this paper in the context of Kac's stochastic framework.

More precisely, we introduce an adaptation of Kac's original stochastic N-particle model that simultaneously includes interactions up to order $M \in \mathbb{N}$ and prove that (1.1) emerges in the $N \to \infty$ limit. The model we propose is motivated by the work of Bobylev, Cercignani, and Gamba [4] on well-posedness and self-similar solutions of an equation that incorporates higher-order collisions between Maxwell molecules. Inspired by [23, 2] we use hierarchy methods to obtain convergence from a certain finite hierarchy of equations to the infinite hierarchy associated to the generalized Kac model. Propagation of chaos then follows as a corollary.

1.1. Higher-order collisions. Let us now introduce higher-order collisions. We shall not specify a concrete transformation map between pre- and postcollisional velocities, but rather work in a general setting that satisfies three conditions, given in the Hypothesis below. We present some examples in section 3.

The transformation law. For every $K = 1, \dots, M$, we assume that we are given a measurable space $\mathbb{S}_K$ with a probability measure b_K, together with a measurable map

$$T_K : \mathbb{S}_K \times \mathbb{R}^{dK} \to \mathbb{R}^{dK} \ .$$

We call K the *order* of the collision, $\mathbb{S}_K$ the space of *scattering angles, b_K* a *collision kernel*, and T_K the *transformation law.*

Throughout this work, we assume the following Hypothesis to be satisfied. Let us denote by S_K the group of permutations of K elements. We will abuse notation and use the same symbol to denote a permutation $\sigma \in S_K$ and its action over the space $\mathbb{R}^{dK}$. Namely, σ stands for the function defined by $\sigma(V) \equiv (V_{\sigma(1)}, \dots, V_{\sigma(K)})$ for $V = (V_1, \dots, V_K) \in \mathbb{R}^{dK}$.

HYPOTHESIS. *For all* $\omega \in \mathbb{S}_K$*, the map* $T_K^\omega \equiv T_K(\omega,\cdot) : \mathbb{R}^{dK} \to \mathbb{R}^{dK}$ *is linear. Additionally, the following hold:*

(H1) T_K^ω *is an isometry.*

(H2) *For all* $\varphi \in C(\mathbb{R}^{dK})$*, it holds that*

$$\int_{\mathbb{S}_K} \varphi\big[(T_K^\omega)^{-1} V\big] \mathrm{d}b_K(\omega) = \int_{\mathbb{S}_K} \varphi[T_K^\omega V] \mathrm{d}b_K(\omega) \qquad \forall V \in \mathbb{R}^{dK} . \tag{1.2}$$

(H3) *For all* $\sigma \in S_K$ *and* $\varphi \in C(\mathbb{R}^{dK})$*, it holds that*

$$\int_{\mathbb{S}_K} \varphi\big[(\sigma \circ T_K^\omega \circ \sigma^{-1})\, V\big] \mathrm{d}b_K(\omega) = \int_{\mathbb{S}_K} \varphi\big[T_K^\omega V\big] \mathrm{d}b_K(\omega) \qquad \forall V \in \mathbb{R}^{dK} . \tag{1.3}$$

The conditions introduced above arise when considering elastic collisions between particles whose pre- and postcollisional velocities are related by the formula

$$(v_1, \ldots, v_K) \longmapsto (v_1^*, \ldots, v_K^*) \equiv T_K^\omega(v_1, \ldots, v_K)\,, \tag{1.4}$$

where $\omega \in \mathbb{S}_K$ is a parameter that labels the directions in which the particles interchange momentum. With this interpretation in mind, we can give physical relevance to the above hypotheses. **(H1)** states that there is *conservation of kinetic energy.* **(H2)** states that, up to an average over the set of scattering angles, the transformation law T_K is an *involution.* **(H3)** states that, up to an average over the set of scattering angles, the transformation law T_K does not depend on the labeling of the particles, e.g., there is no preferred order in which the particles can enter a collision.

Remarks 1.1. A few comments are in order regarding **(H1)**, **(H2)**, and **(H3)**.

(i) Even though our methods can be adapted to include transformation laws that do not satisfy **(H3)**, we include it to make the exposition simpler. Similar assumptions have previously been made in the literature; see, for instance, Definition 2.1(iv) in [8] for an example in the context of the quantum Kac model.

(ii) From a mathematical point of view, we include $K = 1$ since it presents no additional difficulties. Physically, it does not correspond to collisions between the particles but can be understood as an interaction between a single particle and its medium; a famous example is the thermostat model [7].

(iii) In order to accommodate certain models, we *do not* require conservation of momentum to hold:

$$\sum_{i=1}^K v_i^* = \sum_{i=1}^K v_i \ . \tag{1.5}$$

For instance, the above-mentioned thermostat model is such an example. We refer the reader to section 3 for details.

The collisional operators. In this setting, the transformation law T_K defines the collisional operators $Q_K : \prod_{i=1}^K L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ present in the Boltzmann-type equation (1.1) that we derive. More precisely, for $K \geq 2$ these operators are of the form

$$\begin{aligned} &Q_K(f_1, \ldots, f_K)(v_1) \\ &\quad := K \int_{\mathbb{S}_K \times \mathbb{R}^{d(K-1)}} \Big((\otimes_{\ell=1}^K f_\ell)(T_K^\omega V) - (\otimes_{\ell=1}^K f_\ell)(V) \Big) \mathrm{d}b_K(\omega) dv_2 \ldots dv_K \end{aligned} \tag{1.6}$$

and analogously for $K=1$, with $\mathbb{S}_K \times \mathbb{R}^{d(K-1)}$ being replaced by $\mathbb{S}_1$. Notice above that the kernel of Q_K is independent of the *relative velocities* and is *integrable* with respect to the scattering angles. In the context of kinetic theory, such a model can be interpreted as a gas of Maxwell molecules with an angular cut-off.

Extension of the transformation law to N particles. Let $K \in \{1, \dots, M\}$ be a fixed order of collision. We would like to define collisions of order K that happen in a system of N particles; their velocities will be recorded by the so-called *master vector* $V = (v_1, \dots, v_N) \in \mathbb{R}^{dN}$. In order to *select* the particles that undergo a collision, let us denote by

$$\mathcal{I}(K) := \big\{ (i_1, \dots, i_K) \in \{1, \dots, N\}^K \ : i_j \neq i_\ell \text{ for } j \neq \ell \big\} \tag{1.7}$$

the set of all pairwise different indices contained in $\{1, \dots, N\}^K$. Note that we do not require the indices to be ordered; i.e., $i_1 < \cdots < i_K$ may not hold.

Next, let us fix a collection of indices $(i_1, \dots, i_K) \in \mathcal{I}(K)$ and consider a permutation $\sigma \in S_N$ satisfying $\sigma(1) = i_1, \dots, \sigma(K) = i_K$. Then we will work extensively with the new linear map

$$T^\omega_{i_1 \cdots i_K} := \sigma \circ (T^\omega_K \times id_{\mathbb{R}^{d(N-K)}}) \circ \sigma^{-1} : \mathbb{R}^{dN} \to \mathbb{R}^{dN}, \qquad \omega \in \mathbb{S}_K. \tag{1.8}$$

In words, the map $T^\omega_{i_1 \cdots i_K}$ selects the particles labeled by indices $(i_1, \dots, i_K)$ and updates their velocities according to the transformation law (1.4), i.e., $(v_{i_1}, \dots, v_{i_K}) \mapsto (v^*_{i_1}, \dots, v^*_{i_K})$, while leaving the rest invariant.

Remark 1.1. For the special case in which the indices are ordered, meaning that $i_1 < \cdots < i_K$, one can write for $V = (v_1, \dots, v_N) \in \mathbb{R}^{dN}$ the following:

$$T^\omega_{i_1 \cdots i_K} V = (v_1, \dots, v_{i_1 - 1}, v^*_{i_1}, v_{i_1+1}, \dots, v_{i_K - 1}, v^*_{i_K}, v_{i_K + 1}, \dots, v_n), \tag{1.9}$$

where $(v^*_{i_1}, \dots, v^*_{i_K}) \equiv T^\omega_K (v_{i_1}, \dots, v_{i_K}) \in \mathbb{R}^{dK}$.

1.2. Generalized Kac model. As in Kac's original approach for deriving a binary Boltzmann equation, we shall construct a Markov process describing the N-particle system and study the relevant *master equation* governing its dynamics. Details of this construction can be found in section 4.

Our master equation is then given by

$$\begin{cases} \partial_t f_N = \Omega f_N, \\ f_N(0) = f_{N,0} \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN}), \end{cases} \tag{1.10}$$

where L^1_{sym} stands for the space of L^1 functions, invariant under permutation of their variables. The generator $\Omega : L^1(\mathbb{R}^{dN}) \to L^1(\mathbb{R}^{dN})$ is the bounded linear operator determined by the formula

(1.11)

$$\Omega f = N \sum_{K=1}^M \beta_K \sum_{(i_1, \dots, i_K) \in \mathcal{I}(K)} \frac{1}{K! \binom{N}{K}} \int_{\mathbb{S}_K} \big(f \circ T^\omega_{i_1 \cdots i_K} - f \big) \, \mathrm{d}b_K(\omega), \qquad f \in L^1(\mathbb{R}^{dN});$$

we recall that the normalized coefficients $(\beta_K)_{K=1}^M$ were first introduced in (1.1). We note that, since Ω is a bounded linear operator, the solution of the master equation has regularity $f_N \in C^\infty_t(L^1_V)$.

Remark 1.2. We would like to point out that the equation for Ωf above includes an average over the K particles that interact. Consequently, this operation produces a *symmetrization* of the distribution function f. This should not be confused with the symmetry properties that the transformation law T_K^ω may or may not have. For more details, we refer the reader to the discussion contained in subsection 3.2.2.

1.3. Our results in a nutshell. Let us briefly explain the three main results presented in this paper.

(1) *Abstract convergence of hierarchies.* Let $(X^{(s)})_{s\in\mathbb{N}}$ be a collection of Banach spaces. For each $N\in\mathbb{N}$, we consider the following system of equations:

$$\partial_t f_N^{(s)} = \mathscr{C}_{s,s}^N f_N^{(s)} + \cdots + \mathscr{C}_{s,s+M-1}^N f_N^{(s+M-1)}, \quad s\in\mathbb{N}, \tag{1.12}$$

which we call the *N-hierarchy.* Here, each unknown $f_N^{(s)} : [0,T] \to X^{(s)}$ is a time-dependent, vector-valued quantity, and each operator $\mathscr{C}_{s,s+k}^N : X^{(s+k)} \to X^{(s)}$ is linear and bounded, and its operator norm grows at most linearly with $s\in\mathbb{N}$, uniformly in $N\in\mathbb{N}$. We show that if—in an appropriate sense and under additional mild assumptions—the following limits hold,

$$\lim_{N\to\infty} \mathscr{C}_{s,s+k}^N = \mathscr{C}_{s,s+k}^\infty \quad \text{and} \quad \lim_{N\to\infty} f_N^{(s)}(0) = f^{(s)}(0) \qquad \forall s\in\mathbb{N},\ \forall 0\le k\le M-1,$$

then it also holds for later times $t\in[0,T]$ that

$$\lim_{N\to\infty} f_N^{(s)}(t) = f^{(s)}(t),$$

where the limiting objects satisfy the associated infinite system of equations

$$\partial_t f^{(s)} = \mathscr{C}_{s,s}^\infty f^{(s)} + \cdots + \mathscr{C}_{s,s+M-1}^\infty f^{(s+M-1)}, \qquad s\in\mathbb{N},$$

which we call the *infinite hierarchy.* See Definition 3, Definition 4 and Theorem 2.1 for details.

(2) *BBGKY to Boltzmann hierarchy.* Starting from the solution f_N of the master equation (1.10), we show that its sequence of marginals $f_N^{(s)}$ (defined through a partial trace procedure; see (2.19)) satisfies a finite system of equations of the form (1.12), which we shall refer to as the *BBGKY hierarchy.* Under our assumptions on the transformation law T_K^ω and the kernel $\mathrm{d}b_K(\omega)$, every condition of the abstract convergence result is satisfied. Consequently, we can prove that there is convergence to an infinite hierarchy, which we shall refer to as the *Boltzmann hierarchy.* A precise statement can be found in Theorem 2.2.

(3) *Propagation of chaos.* This result concerns the derivation of the Boltzmann equation (1.1). Namely, we assume that the initial data of the master equation $f_{N,0}\in L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$ is such that its sequence of marginals $(f_{N,0}^{(s)})_{s\in\mathbb{N}}$ converges weakly to a tensor product $(f_0^{\otimes s})_{s\in\mathbb{N}}$ for some $f_0\in L^1(\mathbb{R}^d)$. We then prove that for all $t\ge 0$, it holds in the weak sense that

$$\lim_{N\to\infty} f_N^{(s)}(t,\cdot) = f(t,\cdot)^{\otimes s}, \tag{1.13}$$

where $f(t,v)$ is the solution of the Boltzmann equation (2.28), with initial data f_0. This is the content of Theorem 2.3.

Now we provide some context for our results with respect to applications and previous works.

1. *Why is our transformation law abstract?* We decided to require that the transformation law T_K satisfies the general hypotheses **(H1)**, **(H2)**, and **(H3)**.

This allows us to give several examples of transformation laws satisfying the hypotheses (see section 3). Consequently, Theorems 2.2 and 2.3 apply in each of those cases. This is in contrast with respect to most of previous works in the field (see, e.g., [2, 23, 7]) that typically address specific examples.

2. *Why do we have Theorem* 2.1*?* We also remark that our derivation of the generalized Boltzmann equation (1.1), as formulated in Theorem 2.2, is a consequence of the abstract convergence result stated in Theorem 2.1. The motivation for such a level of generality is two-fold: it allows us to identify the estimates that are sufficient for the convergence process, and it provides efficient and robust notation that can be welcome when treating convergence of hierarchies.
3. *Why do we have Theorem* 2.2, *rather than just Theorem* 2.3*?* In contrast to Kac's original approach [28], we prove propagation of chaos as a consequence of convergence of hierarchies (Theorem 2.1). We are therefore able to handle more general initial data. We do not require tensorized initial data.
4. *What are the functional framework novelties of our approach?* Studying convergence of systems with finitely many particles to systems with infinitely many particles has been successfully implemented in the context of derivation of nonlinear PDEs in many cases. These include works on derivation of the nonlinear Schrodinger equations [21] as well as results on derivation of inhomogeneous Boltzmann equations for hard spheres [23]. In this paper, we study space homogeneous systems of Maxwell molecules with angular cut-off. Consequently, the relevant collision operators are bounded in L^1_v, which dictates our main functional framework; see, e.g., Lemma 7.1 for a well-posedness result. This is in contrast to the spaces that have been used in the derivation of the space inhomogeneous Boltzmann equation for hard spheres, which are $L^\infty_{x,v}$-based with exponential weights [23]. However, at the level of the many-particle hierarchies, we employ similarly to [23] time-dependent exponential weights (see the spaces $\boldsymbol{X_\mu}$ in subsection 2.1). To the authors' best knowledge, this is the first application of L^1_v-based spaces in the context of the derivation of Boltzmann equations.

By completion of this work, the authors became aware of the works by Ueno [38] and Tanaka [35, 36] in the late 1960's. These works, as is the case with works on Kac's model, proved propagation of chaos for Markov processes driven by bounded generators, which include some of our models. Their proofs are based on expansion methods, pioneered by Kac [28] and then further developed by McKean [32]. As mentioned above, our result is more general in the sense that we prove convergence of hierarchies (Theorem 2.2) for more general initial data and obtain propagation of chaos as a corollary. Hence the approaches can be understood as complementing each other in terms of methods as well as the results.

Finally, let us mention that, in the last few decades, the problem of deriving an explicit *convergence rate* for the limits (1.13) has received special attention and many models have been investigated; see, e.g., [33, 34, 24, 16, 17, 18, 14]. Such a rate—besides naturally depending on time $t \in \mathbb{R}$, the number of particles $N \in \mathbb{N}$, and the order of the marginals $s \in \mathbb{N}$—comes at the cost of requiring stronger assumptions on both the initial data of the system and the test functions. Even though our methods do not provide a convergence rate, we ask for minimal assumptions on these objects.

Open problems. Let us give a short list of future directions of investigation we believe are interesting.

1. For the model at hand, we expect that a convergence rate in the limit $N \to \infty$ for the marginals (1.13) can be derived, possibly depending nontrivially on

the total number of interactions M. One approach towards finding such a convergence rate would be to employ the framework developed by Mischler and Mouhot [33]. In this context, it would be interesting to understand the *stability* properties of the PDE (1.1) with respect to Wasserstein metrics, as pioneered by Tanaka [36]; see also the work of Cortez and Fontbona [18] and the references therein for more recent developments.

2. It is well known that the inclusion of interactions in a gas of free particles forces the system to reach statistical equilibrium. For the binary Kac model, convergence to equilibrium can be understood in terms of the L^2-*spectral gap* of the (binary) linear operator $\Omega = \Omega(N)$ that depends on particle number; see, e.g., the work of Carlen, Carvalho, and Loss [11]. Hence, it is reasonable to conjecture that the inclusion of additional collisions in the system would *enhance* the convergence to equilibrium. Estimating the spectral gap and its dependence on the order of collisions M would give insight into such a phenomenon.
3. Recently in [3] it has been shown that the addition of ternary interactions among particles that are already allowed to interact binary can in some instances improve moment properties of the corresponding nonlinear equation. The model (1.1) derived in the paper at hand could provide a relatively simple framework for further investigating the question of propagation and generation of moments in nonlinear kinetic equations with higher-order interactions.

Organization of the paper. In section 2, we give precise statements of our three main results. In section 3, a collection of examples that fit our framework are given, and a few adaptations are mentioned. In section 4, we give the details of the construction of the Markov process that gives rise to the master equation (1.10). Theorem 2.1 is proven in section 5, whereas Theorems 2.2 and 2.3 are proven in section 6. Required well-posedness results are proven in section 7, and we include Appendix A for a review of the theory of Markov processes.

2. Main results.

In this section, we state in detail our three main theorems. The first one is an abstract convergence result; in order to state it, we dedicate the next subsection to the necessary functional analytic spaces. This point of view has the advantage of using only minimal estimates satisfied by collision operators. Convergence then happens naturally under the right assumptions.

2.1. The functional framework.

Let $\{X^{(s)}\}_{s\in\mathbb{N}}$ be a collection of Banach spaces, and consider their direct sum

$$X := \bigoplus_{s\in\mathbb{N}} X^{(s)}. \tag{2.1}$$

For any given $\mu \in \mathbb{R}$, we consider the subspace of exponentially weighted sequences

$$X_\mu := \big\{F = (f^{(s)})_{s\in\mathbb{N}} \in X \ : \ \|F\|_\mu := \sup_{s\in\mathbb{N}} e^{\mu s}\|f^{(s)}\|_{X^{(s)}} < \infty\big\}. \tag{2.2}$$

We introduce time dependence as follows. Fix $T > 0$, and consider the weight function

$$\boldsymbol{\mu} : [0,T] \to \mathbb{R}\,, \qquad \boldsymbol{\mu}(t) := -t/T\,. \tag{2.3}$$

We define the Banach space of uniformly bounded, time-dependent sequences as

$$\boldsymbol{X_\mu} := \{\boldsymbol{F} : [0,T] \to X \ : \ \|\boldsymbol{F}\|_{\boldsymbol{\mu}} := \sup_{t\in[0,T]} \|\boldsymbol{F}(t)\|_{\boldsymbol{\mu}(t)} < \infty\}. \tag{2.4}$$

2.1.1. The N-hierarchy. In this subsection, we introduce an abstract version of the BBGKY hierarchy usually found in models that arise from kinetic theory. Before we describe this in detail, let us introduce the following convenient notation: given $K=1,\dots,M$, we work interchangeably with the *lower case* quantities m and k defined through

$$M=m+1 \qquad \text{and} \qquad K=k+1. \tag{2.5}$$

We assume that for $N\geq M$ and $s\in\mathbb{N}$ we are given a collection of bounded linear transformations

$$\begin{gathered}\mathscr{C}^N_{s,s}: X^{(s)} \longrightarrow X^{(s)} \\ \vdots \\ \mathscr{C}^N_{s,s+m}: X^{(s+m)} \longrightarrow X^{(s)},\end{gathered}$$

which we refer to as the *N-hierarchy* operators. Intuitively, one may think of these operators as the components of an infinite matrix, whose entries are nonzero only within a distance m above the diagonal. Let us note that, in this framework, the collection of operators $\{\mathcal{C}^N_{s,s+k}\}_{s=1,k=0}^{\infty,m}$ may be infinite. In applications, however, they are usually a finite collection; see, for example, Remark 6.2.

To the collection of operators mentioned above, we associate the following system of equations, from now on referred to as the *N-hierarchy:*

$$\begin{cases}\partial_t f_N^{(s)} &= \mathscr{C}^N_{s,s} f_N^{(s)} + \cdots + \mathscr{C}^N_{s,s+m} f_N^{(s+m)}, \\ f_N^{(s)}(0) &= f_{N,0}^{(s)} \in X^{(s)},\end{cases} \qquad s\in\mathbb{N}. \tag{2.6}$$

As will become clear during the proof of Theorem 2.1, it will be convenient to write the N-hierarchy in mild form. To this end, we consider the linear operator $\mathscr{C}^N : X\to X$ defined for $F=(f^{(s)})_{s\in\mathbb{N}}$ as

$$(\mathscr{C}^N F)^{(s)} := \mathscr{C}^N_{s,s} f^{(s)} + \cdots + \mathscr{C}^N_{s,s+m} f^{(s+m)} . \tag{2.7}$$

DEFINITION 1. *We say that $\boldsymbol{F}_N = (f_N^{(s)})_{s\in\mathbb{N}} \in \boldsymbol{X_\mu}$ is a mild solution to the N-hierarchy* (2.6) *with initial condition $F_{N,0}=(f_{N,0}^{(s)})_{s\in\mathbb{N}}\in X_0$ if*

$$\boldsymbol{F}_N(t) = F_{N,0} + \int_0^t \mathscr{C}^N \boldsymbol{F}_N(\tau)\,\mathrm{d}\tau \qquad \forall t\in[0,T]. \tag{2.8}$$

We will work under the following assumption. We remind the reader that $m=M-1$ is a fixed natural number.

CONDITION 1. *There exist constants $\{R_k\}_{k=0}^m$, independent of s and N, such that for all $k=0,\dots,m$ there holds that*

$$\|\mathscr{C}^N_{s,s+k} f^{(s+k)}\|_{X^{(s)}} \ \leq\ R_k\, s\, \|f^{(s+k)}\|_{X^{(s+k)}}, \qquad f^{(s+k)}\in X^{(s+k)}. \tag{2.9}$$

Under Condition 1, the following well-posedness result holds. A proof can be found in section 7.

PROPOSITION 2.1. *Assume that the N-hierarchy operators satisfy Condition* 1. *Then, for all $T<(\sum R_k e^k)^{-1}$ and $F_{N,0}\in X_0$, there is a unique mild solution $\boldsymbol{F}_N\in \boldsymbol{X_\mu}$ to the N-hierarchy* (2.6). *In addition, it holds that*

$$\|\boldsymbol{F}_N\|_{\boldsymbol{\mu}} \leq (1-\theta_1)^{-1}\|F_{N,0}\|_0, \qquad \textit{with} \qquad \theta_1 = T\sum_{k=0}^m R_k e^k \in (0,1)\,. \tag{2.10}$$

2.1.2. The infinite hierarchy. If the N-hierarchy operators admit a formal limit when $N \to \infty$, we would like to understand the solutions of the infinite hierarchy they generate. To this end, for each $s \in \mathbb{N}$ we will consider a collection of bounded linear transformations

$$\begin{aligned} \mathscr{C}^{\infty}_{s,s} &: X^{(s)} \longrightarrow X^{(s)} \\ &\vdots \\ \mathscr{C}^{\infty}_{s,s+m} &: X^{(s+m)} \longrightarrow X^{(s)}, \end{aligned}$$

which we call the *infinite hierarchy operators.*

To these operators we associate the *infinite hierarchy,* defined as the infinite system of equations given by

$$\begin{cases} \partial_t f^{(s)} = \mathscr{C}^{\infty}_{s,s} f^{(s)} + \cdots + \mathscr{C}^{\infty}_{s,s+m} f^{(s+m)}, \\ f^{(s)}(0) = f_0^{(s)} \in X^{(s)}, \end{cases} \qquad s \in \mathbb{N}. \tag{2.11}$$

The mild form of the infinite hierarchy is defined analogously. Namely, we consider the linear operator $\mathscr{C}^{\infty} : X \to X$ defined for $F = (f^{(s)})_{s\in\mathbb{N}}$ as

$$(\mathscr{C}^{\infty} F)^{(s)} := \mathscr{C}^{\infty}_{s,s} f^{(s)} + \cdots + \mathscr{C}^{\infty}_{s,s+m} f^{(s+m)}. \tag{2.12}$$

DEFINITION 2. *We say that* $\boldsymbol{F} = (f^{(s)})_{s\in\mathbb{N}} \in \boldsymbol{X}_{\boldsymbol{\mu}}$ *is a mild solution to the infinite hierarchy* (2.11) *with initial condition* $F_0 = (f_0^{(s)})_{s\in\mathbb{N}} \in X_0$ *if*

$$\boldsymbol{F}(t) = F_0 + \int_0^t \mathscr{C}^{\infty} \boldsymbol{F}(\tau)\, \mathrm{d}\tau, \qquad t \in [0,T]. \tag{2.13}$$

We shall assume that the infinite hierarchy operators satisfy an estimate analogous to the one introduced in Condition 1. Namely, we have the following.

CONDITION 2. *There exist constants* $\{\rho_k\}_{k=0}^m$ *such that for all* $k = 0, \dots, m$ *and for all* $s \in \mathbb{N}$

$$\|\mathscr{C}^{\infty}_{s,s+k} f^{(s+k)}\|_{X^{(s)}} \le \rho_k\, s\, \|f^{(s+k)}\|_{X^{(s+k)}}, \qquad f^{(s+k)} \in X^{(s+k)}. \tag{2.14}$$

The following well-posedness result is then available. A proof can be found in section 7.

PROPOSITION 2.2. *Assume that the infinite hierarchy operator* $\mathscr{C}^{\infty}$ *given in* (2.12) *satisfies Condition* 2. *Then, for all* $T < (\sum \rho_k e^k)^{-1}$ *and* $F_0 \in X_0$, *there is a unique mild solution* $\boldsymbol{F} \in \boldsymbol{X}_{\boldsymbol{\mu}}$ *to the infinite hierarchy* (2.11). *In addition, it holds that*

$$\|\boldsymbol{F}\|_{\boldsymbol{\mu}} \le (1-\theta_2)^{-1} \|F_0\|_0, \qquad \text{with} \qquad \theta_2 = T \sum_{k=0}^m \rho_k e^k \in (0,1). \tag{2.15}$$

Remark 2.1. For the rest of the article, we ask the time interval $[0,T]$ to satisfy the following condition:

$$T < m^{-1} T_* \quad \text{with} \quad T_* := \min \left\{ \left(\sum_{k=0}^m R_k e^k \right)^{-1}, \left(\sum_{k=0}^m \rho_k e^k \right)^{-1} \right\}. \tag{2.16}$$

Here, T_* stands for the maximal time for which we can prove simultaneous well-posedness of the two hierarchies; see Propositions 2.1 and 2.2. In particular, T_* is independent of the initial conditions. Consequently, an iteration procedure for proving convergence for all $t \in \mathbb{R}$ is possible, provided global a priori bounds are satisfied by the solutions of the finite and infinite hierarchies, respectively. For the Kac model, these bounds follow from the fact that the solution of the master equation is the density of a probability measure. The extra factor $1/m$ will be used to ensure that certain integral remainder terms converge to zero. See section 5 for details.

2.2. Convergence of hierarchies. The notion of convergence that we are going to study is known in the literature as *convergence of observables.* Before we describe it, we introduce some notation. The bracket $\langle \cdot, \cdot \rangle$ stands for the pairing between $X^{(s)}$ and its dual $X^{(s)*} \equiv (X^{(s)})^*$.

Definition 3. *We introduce the two following notions of convergence:*

1. *The sequence* $(F_N)_{N=1}^{\infty} \in X$ *converges pointwise weakly to* $F \in X$, *abbreviated* $F_N \xrightarrow{pw} F$, *if*

$$\lim_{N\to\infty} \langle f_N^{(s)}, \varphi \rangle = \langle f^{(s)}, \varphi \rangle \qquad \forall s \in \mathbb{N},\ \forall \varphi \in X^{(s)*}, \tag{2.17}$$

where $F_N = (f_N^{(s)})_{s\in\mathbb{N}}$ *and* $F = (f^{(s)})_{s\in\mathbb{N}}$.

2. *The sequence* $\boldsymbol{F}_N : [0,T] \to X$ *converges in observables to* $\boldsymbol{F} : [0,T] \to X$ *if for any* $s \in \mathbb{N}$ *and any* $\varphi \in X^{(s)^*}$

$$\lim_{N\to\infty} \left\langle \boldsymbol{f}_N^{(s)}(t), \varphi \right\rangle = \left\langle \boldsymbol{f}^{(s)}(t), \varphi \right\rangle, \tag{2.18}$$

uniformly in $t \in [0,T]$, *where* $\boldsymbol{F}_N = (\boldsymbol{f}_N^{(s)})_{s\in\mathbb{N}}$ *and* $\boldsymbol{F} = (\boldsymbol{f}^{(s)})_{s\in\mathbb{N}}$.

Let us now make precise the notion in which we understand convergence from the N-hierarchy operators $\mathscr{C}^N$ to the infinite hierarchy operators $\mathscr{C}^\infty$.

Definition 4. *Let* X *be the space introduced in* (2.1). *We say that a sequence of operators* $T^N : X \to X$ *converges to* $T : X \to X$ *if for any sequence* $F_N \in X$ *such that* $F_N \xrightarrow{pw} F$ *it holds true that* $T^N F_N \xrightarrow{pw} TF$.

The following result is our first main theorem; it gives conditions under which convergence in observables occurs, from the finite to the infinite hierarchy.

Theorem 2.1 (convergence of hierarchies). *Assume that the* N*-hierarchy operators* $\mathscr{C}^N$ *satisfy Condition* 1 *and that the infinite hierarchy operators* $\mathscr{C}^\infty$ *satisfy Condition* 2. *Let* $\boldsymbol{F}_N \in \boldsymbol{X}_{\boldsymbol{\mu}}$ *be a mild solution, corresponding to initial data* $F_{N,0} \in X_0$, *of the* N*-hierarchy* (2.8), *and let* $\boldsymbol{F} \in \boldsymbol{X}_{\boldsymbol{\mu}}$ *be a mild solution, corresponding to initial data* $F_0 \in X_0$, *of the infinite hierarchy* (2.13). *In addition, assume that*

(A1) $F_{N,0} \xrightarrow{pw} F_0$,
(A2) $\sup_{N\geq 1} \|F_{N,0}\|_0 < \infty$, *and*
(A3) $\mathscr{C}^N$ *converges to* $\mathscr{C}^\infty$ *in the sense of Definition* 4.

Then $\boldsymbol{F}_N$ *converges in observables to* $\boldsymbol{F}$.

2.3. BBGKY and Boltzmann hierarchies. The next result of this paper concerns the application of Theorem 2.1 to our generalization of the Kac model. In order to state it, let us first introduce the marginals of the solution of the master equation (1.10). Indeed, we consider the following trace map:

$$\mathrm{Tr}_{s+1,\dots,N} : L^1_{\mathrm{sym}}(\mathbb{R}^{dN}) \to L^1_{\mathrm{sym}}(\mathbb{R}^{ds}), \qquad s \in \mathbb{N}, \tag{2.19}$$

where we recall that L^1_{sym} stands for L^1 functions invariant under permutation of their variables. The trace map is then defined for $f \in L^1_{\mathrm{sym}}(\mathbb{R}^d)$ as

$$\operatorname{Tr}_{s+1,\dots,N}[f](V_s) = \begin{cases} \int_{\mathbb{R}^{d(N-s)}} f(V_s, v_{s+1}, \dots, v_N)\, \mathrm{d}v_{s+1}\cdots \mathrm{d}v_N, & s<N, \\ f(V), & s=N, \\ 0, & s>N, \end{cases} \qquad V_s \in \mathbb{R}^{ds} . \tag{2.20}$$

In particular, note that the trace map preserves permutational symmetry.

Let f_N be the solution of the master equation (1.10). We now introduce its marginals as the sequence of functions

$$f_N^{(s)} := \operatorname{Tr}_{s+1,\dots,N}[f_N], \qquad s\in\mathbb{N}. \tag{2.21}$$

One may show that the dynamics of the sequence of sth marginals fits the abstract functional framework introduced above. Namely, by letting $X^{(s)} = L^1_{\mathrm{sym}}(\mathbb{R}^{ds})$ we will show in section 6 that $f_N^{(s)}$ satisfies

$$\partial_t f_N^{(s)} = \mathcal{C}^N_{s,s} f_N^{(s)} + \cdots + \mathcal{C}^N_{s,s+m} f_N^{(s+m)} \qquad \forall s\in\mathbb{N}, \tag{2.22}$$

where $\mathcal{C}^N_{s,s+k} : L^1_{\mathrm{sym}}(\mathbb{R}^{d(s+k)}) \to L^1_{\mathrm{sym}}(\mathbb{R}^{ds})$ are operators that can be computed explicitly. We shall refer to (2.22) as the *BBGKY hierarchy.*

In order to display the structure of the operators $\{\mathcal{C}^N_{s,s+k}\}_{s=1,k=0}^{\infty,m}$, let us first introduce some notation that will be used for the rest of the article.

Notation. Let $s\in\mathbb{N}$ and $k\in\{0,\dots,m\}$. Given $V_s = (v_1,\dots,v_s)\in\mathbb{R}^{ds}$, $v_{s+1},\dots,$ $v_{s+k}\in\mathbb{R}^d$, an index $i\in\{1,\dots,s\}$, and a scattering angle $\omega\in\mathbb{S}_K$, we record the *pre-* and *postcollisional velocities* by the following vectors in $\mathbb{R}^{d(s+k)}$:

$$V_{s+k} := (V_s\,;v_{s+1},\dots,v_{s+k}), \tag{2.23}$$

$$V^{*i}_{s+k} := (v_1,\dots,v_i^*,\dots,v_s\,;v^*_{s+1},\dots,v^*_{s+k}), \tag{2.24}$$

where $(v_i^*, v^*_{s+1},\dots,v^*_{s+k}) \equiv T_K^\omega(v_i, v_{s+1},\dots,v^*_{s+1}) \in \mathbb{R}^{dK}$.

The operators that drive the BBGKY hierarchy then take the form (recall that $K=k+1$)

$$\begin{aligned}(\mathcal{C}^N_{s,s+k} f^{(s+k)})(V_s) &= \frac{\beta_K N}{\binom{N}{K}}\binom{N-s}{K-1}\sum_{i=1}^s \int_{\mathbb{S}_K\times\mathbb{R}^{dk}} \Big(f^{(s+k)}(V^{*i}_{s+k}) - f^{(s+k)}(V_{s+k})\Big)\mathrm{d}b_K(\omega)\mathrm{d}v_{s+1}\dots\mathrm{d}v_{s+k} \\ &\quad + \mathcal{R}^N_{s,s+k}.\end{aligned} \tag{2.25}$$

The operator $\mathcal{R}^N_{s,s+k}$ is a reminder term defined in (6.49) and whose explicit form we do not display here. Importantly, we will also show in section 6 that the operators $\mathcal{C}^N_{s,s+k}$ satisfy Condition 1.

In section 6, we will show that the operators $\mathcal{C}^N_{s,s+k}$ given by (2.25) converge as $N\to\infty$ to the operators $\mathcal{C}^\infty_{s,s+k} : L^1_{\mathrm{sym}}(\mathbb{R}^{d(s+k)}) \to L^1_{\mathrm{sym}}(\mathbb{R}^{ds})$ given by

$$(\mathcal{C}^\infty_{s,s+k} f^{(s+k)})(V_s) = \beta_K K \sum_{i=1}^s \int_{\mathbb{S}_K\times\mathbb{R}^{dk}} \Big(f^{(s+k)}(V^{*i}_{s+k}) - f^{(s+k)}(V_{s+k})\Big)\mathrm{d}b_K(\omega)\mathrm{d}v_{s+1}\dots\mathrm{d}v_{s+k}, \tag{2.26}$$

where V_{s+k} and V_{s+k}^{*i} are as in (2.23) and (2.24), respectively. We verify that these operators satisfy Condition 2 (see Lemma 6.4) and therefore fit the abstract functional framework.

We are now ready to introduce the *Boltzmann hierarchy* as the infinite hierarchy (2.11) with the operators $\mathscr{C}_{s,s+k}^{\infty}$ given by (2.26):

$$\partial_t f^{(s)} = \mathcal{C}_{s,s}^{\infty} f^{(s)} + \cdots + \mathcal{C}_{s,s+m}^{\infty} f^{(s+m)}. \tag{2.27}$$

Our main result concerns the limit from the BBGKY to the Boltzmann hierarchy.

THEOREM 2.2 (from BBGKY to Boltzmann). *Let $X^{(s)} = L^1_{\mathrm{sym}}(\mathbb{R}^{ds})$. Let $\boldsymbol{F}_N$ and $\boldsymbol{F}$ be mild solutions to the BBGKY hierarchy* (2.22) *and Boltzmann hierarchy* (2.27)*, with initial data $F_{N,0} \in X_0$ and $F_0 \in X_0$, respectively. Additionally, assume that $F_{N,0} \xrightarrow{pw} F_0$ and that $\sup_{N\geq 1} \|F_{N,0}\|_0 < \infty$. Then $\boldsymbol{F}_N$ converges in observables to $\boldsymbol{F}$.*

We prove Theorem 2.2 as a corollary of Theorem 2.1; its proof can be found in section 6.

2.4. The Boltzmann equation. We start this subsection by noting that the ansatz $(f^{\otimes s})_{s\in\mathbb{N}}$ is a solution of the Boltzmann hierarchy (2.27) if $f \in C([0,T]; L^1(\mathbb{R}^d))$ solves the following nonlinear equation:

$$\begin{cases} \partial_t f = \beta_1 Q_1(f) + \cdots + \beta_M Q_M(f,\ldots,f), \\ f(0,\cdot) = f_0 \in L^1(\mathbb{R}^d), \end{cases} \tag{2.28}$$

where the collision operators $Q_K : L^1(\mathbb{R}^d)^K \to L^1(\mathbb{R}^d)$ were defined in the introduction; see (1.6). We use the following notion of mild solution.

DEFINITION 5. *We say that $f \in C([0,T]; L^1(\mathbb{R}^d))$ is a mild solution of* (2.28) *corresponding to the initial condition $f_0 \in L^1(\mathbb{R}^d)$ if*

$$f(t) = f_0 + \int_0^t \sum_{K=1}^{M} \beta_K Q_K\Big(f(s),\ldots,f(s)\Big)\mathrm{d}s \qquad \forall t \in [0,T]. \tag{2.29}$$

Global well-posedness for (2.28) was studied in [4] in a slightly different setting. In section 7, we adapt their proof to our situation and obtain the following result.

PROPOSITION 2.3 (global well-posedness). *For all $f_0 \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} f_0(v)\mathrm{d}v = 1$ and $\|f_0\|_{L^1} \leq 1$, there is a unique mild solution $f \in C(\mathbb{R}, L^1(\mathbb{R}^d))$ to the Boltzmann equation* (2.28)*. In addition, $\int_{\mathbb{R}^d} f(t,v)\mathrm{d}v = 1$ and $\|f(t)\|_{L^1} \leq 1$ for all $t \in \mathbb{R}$.*

Remark 2.2. Note that the operators Q_K are continuous in L^1 (see, for instance, Lemma 7.1). Therefore the map $t \mapsto Q_K(f(t),\ldots,f(t)) \in L^1$ is continuous and the fundamental theorem of calculus shows that the global mild solution f of Proposition 2.3 given by (2.29) is of class C^1.

Now we are ready to state our result concerning propagation of chaos for the master equation (1.10). Namely, we prove the following result.

THEOREM 2.3 (propagation of chaos). *Let $f_{N,0} \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$ be nonnegative and normalized to unity:*

$$\|f_{N,0}\|_{L^1} = \int_{\mathbb{R}^{dN}} f_{N,0}(V)dV = 1.$$

Further, assume that its sequence of marginals $(f_{N,0}^{(s)})_{s\in\mathbb{N}}$ converges pointwise weakly to the tensor product $(f_0^{\otimes s})_{s\in\mathbb{N}}$ for some $f_0 \in L^1(\mathbb{R}^d)$. Let $f_N(t)$ be the solution of the master equation (1.10)*, with initial data $f_{N,0}$. Then, for all $t \geq 0$, $s \in \mathbb{N}$, and $\varphi_s \in L^\infty(\mathbb{R}^{ds})$ it holds that*

$$\lim_{N\to\infty} \langle f_N^{(s)}(t,\cdot), \varphi_s\rangle = \left\langle f(t,\cdot)^{\otimes s}, \varphi_s \right\rangle, \tag{2.30}$$

where $f(t,v)$ is the solution of the Boltzmann equation (2.28)*, with initial data f_0.*

Remark 2.3. Since the solution of the master equation $f_N(t)$ is the probability density function of a probability measure (see section 4), it holds that

$$\|f_N(t)\|_{L^1} = \int_{\mathbb{R}^{dN}} f_N(t,V)dV = 1 \qquad \forall t \in \mathbb{R} \tag{2.31}$$

and similarly for its sequence of marginals $f_N^{(s)}(t)$.

3. Applications. In this section, we describe a set of examples that fit the framework introduced in section 1 and further developed in section 2. Namely, they satisfy **(H1)**, **(H2)**, and **(H3)**, and Theorems 2.2 and 2.3 can be applied to each of those models. Some of the examples we consider have already been studied in the literature, and we recover existing results (see examples (1) and (2) below). Example (3), on the other hand, is new.

The following formula is helpful when trying to verify the symmetric condition **(H3)**. Let us regard a linear map $T : \mathbb{R}^{dK} \to \mathbb{R}^{dK}$ as a collection of blocks $T = [T_{ij}]_{i,j=1}^K$, where each $T_{ij} : \mathbb{R}^d \to \mathbb{R}^d$ is linear. Then it holds that

$$\left(\sigma \circ T \circ \sigma^{-1}\right)_{i,j} = T_{\sigma(i),\sigma(j)}, \qquad i,j = 1,\ldots,K \ \ \sigma \in S_K \ . \tag{3.1}$$

3.1. Examples. (1) *Binary collisions.* Let $K = 2$, and take $\mathbb{S}_K = \mathbb{S}_1^{d-1}$, the $(d-1)$-dimensional unit sphere. The transformation law T_B is then defined according to the formulae

$$v_1^* = v_1 + \langle \omega, v_2 - v_1\rangle\, \omega, \tag{3.2}$$

$$v_2^* = v_2 - \langle \omega, v_2 - v_1\rangle\, \omega \tag{3.3}$$

for $\omega \in \mathbb{S}_1^{d-1}$. It is straightforward to verify that T_B is an involution that conserves both energy and momentum. Hence, **(H1)** and **(H2)** are verified. Furthermore, we may write in block form

$$T_B^\omega = \begin{pmatrix} \mathbb{1}_d - \langle\cdot,\omega\rangle\,\omega & \langle\cdot,\omega\rangle\,\omega \\ \langle\cdot,\omega\rangle\,\omega & \mathbb{1}_d - \langle\cdot,\omega\rangle\,\omega \end{pmatrix}, \quad \omega \in \mathbb{S}_1^{d-1}, \tag{3.4}$$

where $\mathbb{1}_d$ is the d-dimensional identity. In particular, it follows that $(T_B^\omega)_{11} = (T_B^\omega)_{22}$ and $(T_B^\omega)_{12} = (T_B^\omega)_{21}$. This observation, combined with (3.1), implies that $\sigma\circ T_B^\omega \circ\sigma^{-1} = T_B^\omega$ for any $\sigma \in S_2$, which in turn implies **(H3)**.

(2) *Kac's toy model.* In dimension $d = 1$, Kac [28] originally considers $\mathbb{S}_2 = (-\pi,\pi)$ and the transformation law $(v_1, v_2) \mapsto T_{toy}^\theta(v_1, v_2)$ determined by the matrix

$$T_{toy}^\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \qquad \text{for } \theta \in (-\pi,\pi). \tag{3.5}$$

Since this is an isometry, it satisfies **(H1)**. We now proceed to verify **(H2)** and **(H3)**. To this end, we calculate that for $\sigma=(1\ 2)\in S_2$ and $\theta\in(-\pi,\pi)$ it holds that

$$\sigma\circ T_{toy}^{\theta}\circ\sigma^{-1} \;=\; [T_{toy}^{\theta}]^{-1} \;=\; \begin{pmatrix}\cos\theta & -\sin\theta\\ +\sin\theta & \cos\theta\end{pmatrix} \;=\; T_{toy}^{-\theta}. \tag{3.6}$$

Consequently, we find that $\sigma\circ T_{toy}^{\theta}\circ\sigma^{-1}=(T_{toy}^{\theta})^{-1}\neq T_{toy}^{\theta}$ for general θ. However, a change of variables $\theta\mapsto-\theta$ shows that **(H2)** and **(H3)** are verified, provided we consider an interaction kernel of the form $\mathrm{d}b_2(\theta)=f(\theta)\mathrm{d}\theta$, where $f\geq 0$ is integrable and even $f(\theta)=f(-\theta)$. These are exactly the conditions considered originally by Kac [28].

(3) *Symmetric collisions of order K.* Consider the set of scattering angles

$$\mathbb{S}_K=\{\omega=(\omega_1,\dots,\omega_K)\in\mathbb{R}^{dK}\mid \omega_1^2+\cdots+\omega_K^2=1\} \tag{3.7}$$

endowed with a probability measure of the form $b(\omega)\mathrm{d}\omega$, where $(b\circ\sigma)(\omega)=b(\omega)$ for all $\sigma\in S_K$. We consider the transformation law T_K given by

$$v_i^*=v_i-2\sum_{\ell=1}^{K}\langle\omega_\ell,v_\ell\rangle\,\omega_i\,,\qquad i\in\{1,\dots,K\}. \tag{3.8}$$

A straightforward calculation shows that T_K is an involution that conserves energy. In addition, the block form representation $[T_K^{\omega}]_{i,j}=\delta_{i,j}\mathbb{1}_d-2\langle\omega_j,\cdot\rangle\,\omega_i$ and (3.1) imply that for all $\sigma\in S_K$ it holds that

$$\sigma\circ T_K^{\omega}\circ\sigma^{-1}=T_K^{\sigma(\omega)},\qquad \omega\in\mathbb{S}_K. \tag{3.9}$$

Since the underlying probability measure is invariant under the change of variables $\omega\mapsto\sigma^{-1}\omega$, one verifies that hypothesis **(H3)** is satisfied. Note that T_K does not conserve momentum. However, if the space $\mathbb{S}_K$ is replaced by

$$\mathbb{S}_K^{'}=\{\omega\in\mathbb{S}_K\,|\,\omega_1+\cdots+\omega_K=0\}, \tag{3.10}$$

one may easily verify that conservation of momentum holds.

3.2. Other models.

In our results, we always assume that **(H1)**, **(H2)**, and **(H3)** are satisfied. We note that there exist models in the literature that fail to satisfy at least one of these conditions and we give two such examples. However, our methods can be adapted to cover theses cases.

3.2.1. Bobylev–Cercignani–Gamba model.

For $K\leq M$, suppose that one is given scalar velocities $(v_1,\dots,v_K)\in\mathbb{R}^K$. In [4], the authors propose a model for economic games in which the particles (or players) undergo a transformation law $T_{a,b}$ of the form

$$v_i^*=av_i+b\sum_{j\neq i}v_j\,,\qquad i\in\{1,\dots,K\}, \tag{3.11}$$

where the real-valued coefficients a and b are random variables on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Note that even when $K=2$, this transformation fails to conserve energy unless the coefficients are heavily constrained; conservation of energy would force $|a^2-b^2|=1$. Note, however, that the relation $[\sigma\circ T_{a,b}\circ\sigma^{-1}]_{i,j}=[T_{a,b}]_{\sigma(i),\sigma(j)}=[T_{a,b}]_{i,j}$ implies that condition **(H3)** is verified, independently of the underlying probability space or the specific structure of the coefficients a and b.

One may still consider the situation in which $d(\omega) := |\det T^{\omega}_{a,b}| > 0$, that is, the case for which $T_{a,b}$ is invertible. By keeping track of the $d(\omega)$ factor, we expect that results analogous to Theorems 2.2 and 2.3 can be proven, leading to a Boltzmann equation (2.28), with a collisional operator given by

$$\begin{aligned} &Q_K(f_1,\dots,f_K)(v_1)\\ &\quad := K\int_{\mathbb{S}_K\times\mathbb{R}^{(K-1)}} \left(\frac{1}{d(\omega)}(\otimes_{\ell=1}^K f_\ell)\big([T_{a,b}]^{-1}V\big) - (\otimes_{\ell=1}^K f_\ell)(V)\right) \mathrm{d}b_K(\omega)dv_2\dots dv_K. \end{aligned}$$

3.2.2. Nonsymmetric ternary collisions. Let us focus on the ternary case $K=3$ and consider the $(2d-1)$-unit sphere $\mathbb{S}_3 := \mathbb{S}^{2d-1}$ with the usual surface measure $\mathrm{d}\omega$. As noted in [1], the relevant transformation law T_{ter} is defined as

$$v_1^* = v_1 - c(v_1,v_2,v_3;\omega)(\omega_1+\omega_2), \tag{3.12}$$

$$v_2^* = v_2 + c(v_1,v_2,v_3;\omega)\,\omega_1, \qquad c(v_1,v_2,v_3;\omega) = \frac{\langle\omega_1, v_2-v_1\rangle + \langle\omega_2, v_3-v_1\rangle}{1+\langle\omega_1,\omega_2\rangle}, \tag{3.13}$$

$$v_3^* = v_3 + c(v_1,v_2,v_3;\omega)\,\omega_2, \tag{3.14}$$

where $\omega = (\omega_1,\omega_2)\in\mathbb{S}^{2d-1}$. Despite conserving energy and momentum, hypothesis **(H3)** is not satisfied for this model. Let us further explain. We note that symmetries can be understood at two levels.

1. *At the level of the collisional laws of the particles.* This is a condition on a map $T:\mathbb{R}^{dK}\to\mathbb{R}^{dK}$ of K variables. In this level, the interaction in [1] is not symmetric since there is one preferred particle. One can check by hand that the ternary interaction from [1] does not satisfy **(H3)**.
2. *At the level of the evolution of the distribution function.* In this level, symmetry of the function is recovered by the "average" over the K variables. However, symmetry at the *particle* level is not: at each interaction, particles still interact asymmetrically.

We expect, however, that our methods can be adapted to show that similar results hold true, leading to a Boltzmann equation with a collisional operator of the form

$$Q_3 = Q_3^{(1)} + 2Q_3^{(2)},$$

where, for $f\in L^1(\mathbb{R})$, we have

$$Q_3^{(1)}[f](v_1) = \int_{\mathbb{S}_3\times\mathbb{R}^2} \Big(f(v_1^*)f(v_2^*)f(v_3^*) - f(v_1)f(v_2)f(v_3)\Big)\mathrm{d}b(\omega)dv_2dv_3, \tag{3.15}$$

$$Q_3^{(2)}[f](v_2) = \int_{\mathbb{S}_3\times\mathbb{R}^2} \Big(f(v_1^*)f(v_2^*)f(v_3^*) - f(v_1)f(v_2)f(v_3)\Big)\mathrm{d}b(\omega)dv_1dv_3. \tag{3.16}$$

Namely, the strategy of the proof will follow along the same lines but would require keeping track of more complex combinatorics.

4. The master equation. In order to accommodate higher-order interactions among particles, in this section we construct a new Markov process. We are inspired by the pioneering work of Kac [28], where the author outlined the procedure for constructing the Markov process corresponding to binary interactions. Our Markov process then leads to the master equation (1.10).

For simplicity of exposition, we work with Euclidean space $\mathbb{R}^{dN}$, instead of restricting ourselves to the *energy spheres*

$$\mathcal{E}_N := \{V = (v_1, \dots, v_N) \in \mathbb{R}^{dN} \;:\; |V| = \sqrt{N}\}. \tag{4.1}$$

Our methods can be easily adapted to incorporate restrictions to $\mathcal{E}_N$ (since conservation of kinetic energy satisfied by the transformation law leaves the energy spheres invariant).

First, we describe the heuristics behind constructing our Markov process. As noted above, we incorporate higher-order collisions given by the transformation law (1.4). Then we give a sketch of the mathematical details of its construction as a jump process. We refer the reader to Appendix A for a brief review of the theory of Markov processes, including the notation that will be extensively used in this section.

In order to construct the continuous time Markov process $\mathbf{V}_N$, we will first construct the simpler discrete time process $\mathbf{Y}_N$, where $\mathbf{Y}_N(n)$ represents the state of our N-particle system after the nth collision. Recall that we fix $(\mathbb{S}_K, T_K, b_K)$ with $K = 1, \dots, M$ as introduced in section 1. Fix positive parameters $\{\beta_K\}_{K=1}^M$ that satisfy the normalization condition

$$\beta_1 + \cdots + \beta_M = 1. \tag{4.2}$$

Here the parameters β_K represent the probability that a given collision will be of order K. Given the distribution of $\mathbf{Y}_N(n)$, we obtain the distribution of $\mathbf{Y}_N(n+1)$, the system after one collision, by following the steps:

1. Select $K \in \{1, \dots, M\}$ with probability β_K. This determines the order of the system's next collision.
2. Select which K of the N particles will undergo this collision by choosing an ordered index $(i_1, \dots, i_K)$ uniformly from $\mathcal{I}_K$. This choice has probability $(K!)^{-1}\binom{N}{K}^{-1}$.
3. Select the impact parameter $\omega \in \mathbb{S}_K$ according to the law $\mathrm{d}b_K(\omega)$.
4. Update the velocities as follows:
$$\mathbf{Y}_N(n+1) = T^{\omega}_{i_1,\dots,i_K}(v_1, \dots, v_N),$$
where $T^{\omega}_{i_1,\dots,i_K}$ is given by (1.8).

If we start with a given initial distribution $\mathbf{Y}_N(0)$ of our N-particle system, we can *formally* construct our process $\mathbf{Y}_N$ completely by repeating the above steps.

To construct $\mathbf{Y}_N$ *rigorously*, we introduce a Markov transition function acting on $V \in \mathbb{R}^{dN}$ and a Borel set $B \in \mathscr{B}(\mathbb{R}^{dN})$

(4.3)

$$\mu_N(V, B) := \sum_{K=1}^{M} \beta_K \sum_{(i_1,\dots,i_K)} \frac{1}{K!\binom{N}{K}} \int_{\mathbb{S}_K} \mathbb{1}_B(T^{\omega}_{i_1,\dots,i_K} V) \mathrm{d}b_K(\omega), \qquad V \in \mathbb{R}^{dN},\ B \in \mathscr{B}(\mathbb{R}^{dN}),$$

whose (bounded) generator $P_N : C_b(\mathbb{R}^{dN}) \to C_b(\mathbb{R}^{dN})$ satisfies

$$\begin{aligned}(P_N\varphi)(V) &:= \int_{\mathbb{R}^{dN}} \varphi(U) \mu_N(V, dU) \\ &= \sum_{K=1}^{M} \beta_K \sum_{(i_1,\dots,i_K)} \frac{1}{K!\binom{N}{K}} \int_{\mathbb{S}_K} \varphi(T^{\omega}_{i_1,\dots,i_K} V)\, \mathrm{d}b_K(\omega) \quad V \in \mathbb{R}^{dN}.\end{aligned} \tag{4.4}$$

Given $f_{N,0} \in \mathrm{Prob}(\mathbb{R}^{dN})$, the space of probability measures on $\mathbb{R}^{dN}$, by Proposition A.2 we can find a probability space $(\Sigma, \mathscr{F}, \mathbb{P})$ and a Markov chain $\{\mathbf{Y}_N(n)\}_{n=0}^{\infty} : \Sigma \times \mathbb{N}_0 \to \mathbb{R}^{dN}$ whose transition function is μ_N and whose initial law is determined by $f_{N,0}$. In other words, it holds that for all $n \in \mathbb{N}_0$ and $B \in \mathscr{B}(\mathbb{R}^{dN})$

$$\begin{aligned} &\mathbb{P}[\mathbf{Y}_N(n+1) \in B | \mathbf{Y}_N(0), \ldots, \mathbf{Y}_N(n)] = \mu_N(\mathbf{Y}_N(n), B), \\ &\mathbb{P}[\mathbf{Y}_N(0) \in B] = f_{N,0}(B). \end{aligned} \tag{4.5}$$

By computing the one step transition probability for $\mathbf{Y}_N$, it can be checked that μ_N given in (4.3) is the correct transition function for our process $\mathbf{Y}_N$.

In order to introduce continuous time into our process, consider an independent Poisson process $\{M(t)\}_{t=0}^{\infty}$ with rate N (see Definition 14 in Appendix A) and define the Markov process $\mathbf{V}_N(t)$ as the *jump process*

$$\mathbf{V}_N(t) := \mathbf{Y}_N(M(t)). \tag{4.6}$$

In particular, it can be shown that this jump process corresponds to the transition semigroup $\{T(t)\}_{t\geq 0}$ whose (bounded) generator is

$$\mathscr{L}_N := N(P_N - \mathrm{id}) : C_b(\mathbb{R}^{dN}) \to C_b(\mathbb{R}^{dN}), \tag{4.7}$$

where P_N is defined in (4.4). The reader is referred to section 2.2 of [22] for details.

Our starting point for the derivation of the Boltzmann equation (1.1) will be the dynamics associated to the law of the process $V_N(t)$. More precisely, let us denote its law by $F_N(t,\cdot)$. This is a probability measure on $\mathbb{R}^{dN}$, invariant under permutations, with the symmetric property being equivalent to the particles being indistinguishable. We make the additional assumption that the initial data has a symmetric density $f_{N,0} \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$. Consequently, F_N has a density f_N that evolves according to the *master equation*

$$\begin{cases} \partial_t f_N = \Omega f_N, \\ f_N(0) = f_{N,0} \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN}), \end{cases} \tag{4.8}$$

where the generator $\Omega : L^1_{\mathrm{sym}}(\mathbb{R}^{dN}) \to L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$ is the bounded linear operator determined by the formula

(4.9)
$$\Omega f = N \sum_{K=1}^{M} \beta_K \sum_{i_1 \cdots i_K} \frac{1}{K!\binom{N}{K}} \int_{\mathbb{S}_K} \left(f \circ T^{\omega}_{i_1,\ldots,i_K} - f \right) \mathrm{d}b_K(\omega), \qquad f \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN}).$$

Remark 4.1 (relationship to the deterministic setting). The Liouville equation is the deterministic analogue of the master equation (4.8). Furthermore, N is chosen for the rate of the Poisson process $M(t)$ in (4.6) to ensure a constant number of collisions per unit time per particle in the limit $N \to \infty$ and is analogous to the Boltzmann–Grad scaling in the deterministic setting.

5. Proof of Theorem 2.1.

Throughout this section, we assume that the estimates contained in Conditions 1 and 2 are satisfied, together with assuming that $T > 0$ satisfies (2.16). First, we introduce some notation and prove some preliminary inequalities.

In what follows, we will be using the same notation introduced in subsection 2.1. For $s\in\mathbb{N}$, let us introduce the canonical projections

$$\pi_s : X = \bigoplus_{r\in\mathbb{N}} X^{(r)} \longrightarrow X^{(s)} \tag{5.1}$$

defined for $F=(f^{(s)})_{s\in\mathbb{N}}\in X$ as $\pi_s(F):=f^{(s)}$. In particular, in terms of the objects $\boldsymbol{F}:[0,T]\to X$ and $\boldsymbol{F}_N:[0,T]\to X$, convergence of observables (see Definition 3) is equivalent to the following statement: for all $s\in\mathbb{N}$ and for all $\varphi_s\in X^{(s)*}$, there holds that

$$\lim_{N\to\infty}\langle \pi_s \boldsymbol{F}_N(t),\varphi_s\rangle = \langle \pi_s\boldsymbol{F}(t),\varphi_s\rangle \tag{5.2}$$

uniformly in $t\in[0,T]$.

Let $\mathscr{C}^N,\mathscr{C}^\infty : X\to X$ be the linear transformations introduced in (2.7) and (2.12), respectively. The introduction of the projections $(\pi_s)_{s\in\mathbb{N}}$ will be particularly useful for proving norm estimates for the sth components of the *iterated* powers of $\mathscr{C}^\infty$ ($\mathscr{C}^N$, resp.), namely for the operators

$$(\mathscr{C}^\infty)^n = \underbrace{\mathscr{C}^\infty\circ\cdots\circ\mathscr{C}^\infty}_{n\text{ times}}, \qquad n\in\mathbb{N}. \tag{5.3}$$

More precisely, the following lemma holds true.

LEMMA 5.1. (a) *If $\mathscr{C}^N$ satisfies Condition 1, then for every $\ell\in\mathbb{N}$, $s\in\mathbb{N}$, and $F\in X$ there holds that*

$$\begin{aligned}\|\pi_s\big[(\mathscr{C}^N)^\ell F\big]\|_{X^{(s)}} \le \sum_{k_1=0}^m\cdots\sum_{k_\ell=0}^m & s(s+k_1)\cdots(s+k_1+\cdots+k_{\ell-1})\\ &\times R_{k_1}\cdots R_{k_\ell}\|\pi_{s+k_1\cdots+k_\ell}F\|_{X^{(s+k_1\cdots+k_\ell)}}\ .\end{aligned} \tag{5.4}$$

(b) *If $\mathscr{C}^\infty$ satisfies Condition 2, then for every $\ell\in\mathbb{N}$, $s\in\mathbb{N}$, and $F\in X$ there holds that*

$$\begin{aligned}\|\pi_s\big[(\mathscr{C}^\infty)^\ell F\big]\|_{X^{(s)}} \le \sum_{k_1=0}^m\cdots\sum_{k_\ell=0}^m & s(s+k_1)\cdots(s+k_1+\cdots+k_{\ell-1})\\ &\times \rho_{k_1}\cdots \rho_{k_\ell}\|\pi_{s+k_1\cdots+k_\ell}F\|_{X^{(s+k_1\cdots+k_\ell)}}.\end{aligned} \tag{5.5}$$

Proof. We shall only present a proof for (b); that of (a) is identical. In what follows, we omit the subscript $X^{(s)}$ from the norms $\|\cdot\|_{X^{(s)}}$. The proof goes by induction on $\ell\in\mathbb{N}$. Indeed, for $\ell=1$ let $s\in\mathbb{N}$, let $F=(f^{(s)})_{s\in\mathbb{N}}\in X$, and estimate using Condition 2 that

$$\begin{aligned}\|\pi_s\big[\mathscr{C}^\infty F\big]\| &= \|\mathscr{C}^\infty_{s,s}f^{(s)}+\cdots+\mathscr{C}^\infty_{s,s+m}f^{(s+m)}\| \le s\big(\rho_0\|f^{(s)}\|+\cdots+\rho_m\|f^{(s+m)}\|\big)\\ &= s\sum\rho_k\|\pi_{s+k}\,F\|\,.\end{aligned} \tag{5.6}$$

Assume now that the result holds up to $\ell\in\mathbb{N}$. Then, for $s\in\mathbb{N}$ and $F\in X$, we have that

$$\begin{aligned}\|\pi_s\big[(\mathscr{C}^\infty)^{\ell+1}F\big]\| \le & \sum_{k_1=0}^{m}\cdots\sum_{k_\ell=0}^{m} s(s+k_1)\cdots(s+k_1+\cdots+k_{\ell-1}) \\ & \times \rho_{k_1}\cdots\rho_{k_\ell}\|\pi_{s+k_1\cdots+k_\ell}\mathscr{C}^\infty F\| \\ \le & \sum_{k_1=0}^{m}\cdots\sum_{k_\ell=0}^{m} s(s+k_1)\cdots(s+k_1+\cdots+k_{\ell-1}) \\ & \times \rho_{k_1}\cdots\rho_{k_\ell}\sum_{k_{\ell+1}=0}^{m}(s+k_1+\cdots k_\ell)\rho_{k_{\ell+1}}\|\pi_{s+k_1+\cdots k_{\ell+1}}F\| \ . \end{aligned} \tag{5.7}$$

This finishes the proof of the lemma. ∎

The following lemma will be useful throughout the proof of convergence. We recall that the well-posedness time T_* was defined in (2.16).

LEMMA 5.2. *Let $s\in\mathbb{N}$, let $\mu\ge -1$, and let $n\ge 10$.*
(a) *If $\mathscr{C}^N$ satisfies Condition 1, then for all $F\in X_\mu$ there holds that*

$$\|\pi_s\big[(\mathscr{C}^N)^n F\big]\|_{X^{(s)}} \ \le \ s e^{-\mu s}\, n!\,(mT_*^{-1})^n\,(e\,n)^{s/m}\|F\|_\mu \ . \tag{5.8}$$

(b) *If $\mathscr{C}^\infty$ satisfies Condition 2, then for all $F\in X_\mu$ there holds that*

$$\|\pi_s\big[(\mathscr{C}^\infty)^n F\big]\|_{X^{(s)}} \ \le \ s e^{-\mu s}\, n!\,(mT_*^{-1})^n\,(e\,n)^{s/m}\|F\|_\mu \ . \tag{5.9}$$

Proof. Similarly as before, we shall only present a proof of (b). Let s,n,μ be as in the statement of the lemma, and for the sake of the proof let us denote $\alpha = s/m$. Then, for any $0\le k_1,\dots,k_n\le m$, we have the following upper bound:

$$\begin{aligned} s(s+k_1)\cdots(s+k_1+\cdots+k_{n-1}) &\le s(s+m)\cdots(s+(n-1)m) \\ &= sm^{n-1}(sm^{-1}+1)\cdots(sm^{-1}+(n-1)) \\ &= sm^{n-1}(\alpha+1)\cdots(\alpha+(n-1)) \\ &= sm^{n-1}(n-1)!(\alpha+1)\cdots\Big(\frac{\alpha}{n-1}+1\Big) \\ &\le s\,m^n n!(\alpha+1)\cdots\Big(\frac{\alpha}{n}+1\Big). \end{aligned} \tag{5.10}$$

For notational convenience, we have replaced $n-1$ by n; since we are only interested in the asymptotic behavior when $n\to\infty$, such a replacement is harmless. Next, using the fact that $\log(1+x)\le x$ for all $x\ge 0$, one finds that

$$\begin{aligned} (\alpha+1)\cdots\Big(\frac{\alpha}{n}+1\Big) &= \exp\log\left((\alpha+1)\cdots\Big(\frac{\alpha}{n}+1\Big)\right) \\ &= \exp\left(\log(\alpha+1)+\cdots+\log\Big(\frac{\alpha}{n}+1\Big)\right) \\ &\le \exp\Big(\alpha\big(1+1/2+\cdots+1/n\big)\Big). \end{aligned} \tag{5.11}$$

For $n\ge 10$, one has the standard bound $\sum_{j=1}^n 1/j \le \log(n)+1$. Consequently, we find that

$$(\alpha+1)\cdots\Big(\frac{\alpha}{n}+1\Big) \le \exp\Big(\alpha\log(n)+\alpha\Big) = (e\,n)^{s/m}. \tag{5.12}$$

Next, we use the definition of the norm $\|\cdot\|_\mu$ (see (2.2)) to find that

$$\|\pi_{s+k_1+\cdots+k_n}F\|_{X^{(s+k_1+\ldots+k_n)}} \le \exp\big(-\mu(s+k_1+\cdots+k_n)\big)\|F\|_\mu. \tag{5.13}$$

Hence, by Lemma 5.1 and (5.10), (5.12), (5.13), we find that

$$\begin{aligned}\|\pi_s(\mathscr{C}^\infty)^n F\|_{X^{(s)}} \ \le\ & s\, m^n n!\,(e\,n)^{s/m}\\ & \times \sum_{k_1\ldots k_n}^{m} \exp\big(-\mu(s+k_1+\cdots+k_{n-1})\big)\rho_{k_1}\cdots\rho_{k_n}\,\|F\|_\mu,\end{aligned} \tag{5.14}$$

from which the desired estimate follows after elementary manipulations, taking into account the definition of the well-posedness time T_*; see (2.16). □

We are now ready to give a proof of Theorem 2.1.

Proof of Theorem 2.1. Let $\boldsymbol{F}_N = (f_N^{(s)})_{s\in\mathbb{N}} \in \boldsymbol{X_\mu}$ and $\boldsymbol{F} = (f^{(s)})_{s\in\mathbb{N}} \in \boldsymbol{X_\mu}$ be as in the statement of Theorem 2.1, with initial data $F_{N,0} = (f_{N,0}^{(s)})_{s\in\mathbb{N}} \in X_0$ and $F_0 = (f_0^{(s)})_{s\in\mathbb{N}} \in X_0$, respectively. Recall that existence and uniqueness of mild solutions of both hierarchies is guaranteed by Propositions 2.1 and 2.2. The idea of the proof is as follows: for fixed $s\in\mathbb{N}$, starting from both the finite and the infinite hierarchies in mild formulation, we iterate the integral formulas $n\in\mathbb{N}$ times. Next, we show that the initial conditions match in the limit $N\to\infty$ and the integral remainder term vanishes as $n\to\infty$, uniformly in N.

Let us be more precise. First, we write the mild formulation of the solutions of both hierarchies:

$$\boldsymbol{F}_N(t) = F_{N,0} + \int_0^t \mathscr{C}^N \boldsymbol{F}_N(\tau)\mathrm{d}\tau, \tag{5.15}$$

$$\boldsymbol{F}(t) = F_0 + \int_0^t \mathscr{C}^\infty \boldsymbol{F}(\tau)\mathrm{d}\tau. \tag{5.16}$$

Next, let us fix $s\in\mathbb{N}$ and iterate n times the above equations to get

$$\boldsymbol{F}_N(t) \ =\ \sum_{\ell=0}^n \frac{t^\ell}{\ell!} x(\mathscr{C}^N)^\ell F_{N,0} \ +\ \int_0^t\cdots\int_0^{t_n} (\mathscr{C}^N)^{n+1}\boldsymbol{F}_N(t_{n+1})\,\mathrm{d}t_{n+1}\cdots\mathrm{d}t_1, \tag{5.17}$$

$$\boldsymbol{F}(t) \ =\ \sum_{\ell=0}^n \frac{t^\ell}{\ell!}(\mathscr{C}^\infty)^\ell F_0 \ +\ \int_0^t\cdots\int_0^{t_n} (\mathscr{C}^\infty)^{n+1}\boldsymbol{F}(t_{n+1})\,\mathrm{d}t_{n+1}\cdots\mathrm{d}t_1. \tag{5.18}$$

Once we project with π_s and consider the pairing with $\varphi_s\in X^{(s)*}$, we note that the contribution to this difference arises due to two terms:

$$\big|\langle \pi_s\boldsymbol{F}_N(t),\varphi_s\rangle - \langle\pi_s\boldsymbol{F}(t),\varphi_s\rangle\big| \ \le\ \mathcal{S}_{N,n}(t) \ +\ \mathcal{I}_{N,n}(t), \tag{5.19}$$

where $\mathcal{S}_{N,n}(t)$ is the *sum* given by

$$\mathcal{S}_{N,n}(t) \ :=\ \sum_{\ell=0}^n \frac{t^\ell}{\ell!}\,\big|\langle\pi_s(\mathscr{C}^N)^\ell F_{N,0},\varphi_s\rangle - \big\langle\pi_s(\mathscr{C}^\infty)^\ell F_0,\varphi_s\big\rangle\big| \tag{5.20}$$

and where $\mathcal{I}_{N,n}(t)$ is an *integral remainder* term defined as

(5.21)

$$\begin{aligned}&\mathcal{I}_{N,n}(t)\\ &:= \int_0^t\cdots\int_0^{t_n}\Big(\|\pi_s(\mathscr{C}^N)^{n+1}\boldsymbol{F}_N(t_{n+1})\|_{X^{(s)}} + \|\pi_s(\mathscr{C}^\infty)^{n+1}\boldsymbol{F}(t_{n+1})\,\|_{X^{(s)}}\Big)\mathrm{d}t_{n+1}\cdots\mathrm{d}t_1,\end{aligned}$$

and we assume without loss of generality that $\|\varphi_s\|_{X^{(s)*}} \leq 1$. We study these two terms separately.

Integral Remainder Terms $\mathcal{I}_{N,n}$. It suffices to estimate the time integrals, with respect to n, uniformly in N. We actually show that each integral separately converges to zero in $X^{(s)}$ norm, once we project via the map π_s. Since the estimates are identical for $\boldsymbol{F}_N(t)$ and $\boldsymbol{F}(t)$, we only present a proof for the latter.

First, we introduce the following notation, convenient for estimating the nested integrals:

$$\mathrm{d}\bar{t}_{n+1} \equiv \mathrm{d}t_{n+1} \cdots \mathrm{d}t_1, \qquad n \in \mathbb{N}.$$

Further, we recall that in section 2 we have introduced the function

$$\boldsymbol{\mu}(t) = -t/T, \qquad t \in [0,T],$$

where $T < m^{-1}T_*$; see (2.16). In view of Lemma 5.2, we find that, for all $n \in \mathbb{N}$ and $t_{n+1} \leq t \leq T$, the following estimate holds:

$$\begin{aligned} \|\pi_s(\mathscr{C}^\infty)^{n+1}\boldsymbol{F}(t_{n+1})\|_{X^{(s)}} \\ \leq s e^{-\boldsymbol{\mu}(t_{n+1})s}\,(n+1)!\,(mT_*^{-1})^{n+1}\,[e(n+1)]^{s/m}\|\boldsymbol{F}(t_{n+1})\|_{\boldsymbol{\mu}(t_{n+1})}. \end{aligned} \tag{5.22}$$

Consequently, we find that

$$\begin{aligned} &\int_0^t \cdots \int_0^{t_n} \|\pi_s(\mathscr{C}^\infty)^{n+1}\boldsymbol{F}(t_{n+1})\|_{X^{(s)}}\mathrm{d}\bar{t}_{n+1} \leq s(n+1)!(mT_*^{-1})^{n+1}\,[e(n+1)]^{s/m} \\ &\qquad \times \int_0^t \cdots \int_0^{t_n} e^{-\boldsymbol{\mu}(t_{n+1})s}\|\boldsymbol{F}(t_{n+1})\|_{\boldsymbol{\mu}(t_{n+1})}\mathrm{d}\bar{t}_{n+1} \\ &\quad \leq s(n+1)!(mT_*^{-1})^{n+1}\,[e(n+1)]^{s/m} \\ &\qquad \times \|\boldsymbol{F}\|_{\boldsymbol{\mu}} \int_0^t \cdots \int_0^{t_n} e^{-\boldsymbol{\mu}(t_{n+1})s}\mathrm{d}\bar{t}_{n+1} \\ &\quad \leq s(n+1)!(mT_*^{-1})^{n+1}\,[e(n+1)]^{s/m} \\ &\qquad \times \|\boldsymbol{F}\|_{\boldsymbol{\mu}} e^s \frac{T^{n+1}}{(n+1)!} \\ &\quad = s(mTT_*^{-1})^{n+1}[e(n+1)]^{s/m}e^s\|\boldsymbol{F}\|_{\boldsymbol{\mu}}. \end{aligned} \tag{5.23}$$

We recall that T was chosen small enough in (2.16) so that $mTT_*^{-1} < 1$ holds true. Therefore, as $n \to \infty$, the integral remainder term vanishes.

Controlling the Sum $\mathcal{S}_{N,n}$. First, we show that the following result holds.

Lemma 5.3. *Let $F_N \in X$ converge pointwise weakly to $F \in X$, and let $\mathscr{C}^N$ converge to $\mathscr{C}^\infty$ in the sense of Definition 4. Then, for all $\ell \in \mathbb{N}$, it holds that $(\mathscr{C}^N)^\ell F_N$ converges pointwise weakly to $(\mathscr{C}^\infty)^\ell F$. In other words, for all $s \in \mathbb{N}$, $\ell \in \mathbb{N}$, and $\varphi_s \in X^{(s)*}$ it holds that*

$$\lim_{N\to\infty} \langle \pi_s\big[(\mathscr{C}^N)^\ell F_N\big], \varphi_s\rangle = \big\langle \pi_s\big[(\mathscr{C}^\infty)^\ell F\big], \varphi_s\big\rangle.$$

Proof. The proof goes by induction on $\ell \in \mathbb{N}$. The case $\ell = 1$ follows from the definition of convergence from $\mathscr{C}^N$ to $\mathscr{C}^\infty$. Assume now that the result holds for $\ell \in \mathbb{N}$, i.e., $G_N = (\mathscr{C}^N)^\ell F_N$ converges weakly to $G = (\mathscr{C}^\infty)^\ell F$. It follows that $\mathscr{C}^N G_N$ converges pointwise weakly to $\mathscr{C}^\infty G$. This finishes the proof. ∎

Conclusion. First, we take the limit $N \to \infty$. Namely we put our two estimates together to find that for all $n \geq 1$ there holds that

$$\begin{aligned} \limsup_{N\to\infty} \big| \langle \pi_s \boldsymbol{F}_N(t), \varphi_s \rangle - \langle \pi_s \boldsymbol{F}(t), \varphi_s \rangle \big| &\leq \limsup_{N\to\infty} \mathcal{S}_{N,n}(t) + \limsup_{N\to\infty} \mathcal{I}_{N,n}(t) \\ &\leq s(mTT_*^{-1})^{n+1}[e(n+1)]^{s/m} e^s \left(\|\boldsymbol{F}\|_{\boldsymbol{\mu}} + \sup_{N\in\mathbb{N}} \|\boldsymbol{F}_N\|_{\boldsymbol{\mu}} \right). \end{aligned} \tag{5.24}$$

Thanks to Proposition 2.1, one has that $\|\boldsymbol{F}_N\|_{\boldsymbol{\mu}} \leq (1-\theta_2)^{-1} \|F_{N,0}\|_0$ for all $N \geq 1$. Thus, $\sup_{N\in\mathbb{N}} \|\boldsymbol{F}_N\|_{\boldsymbol{\mu}} < \infty$ due to our assumptions on the initial data. The conclusion of the theorem now follows after we take the $n \to \infty$ limit. ∎

6. Proofs of Theorems 2.2 and 2.3.

Throughout this section, f_N denotes the solution of the master equation (1.10), and $(f_N^{(s)})_{s\in\mathbb{N}}$ denotes its sequence of marginals, defined in (2.21). We recall that these quantities are symmetric with respect to the permutation of their variables.

6.1. Calculation of BBGKY.

In what follows, we fix the number of particles $N \geq M$ and some order $s \leq N$ of the marginals. We start with the following calculation:

$$\partial_t f_N^{(s)} = \partial_t \mathrm{Tr}_{s+1,\dots,N}\big(f_N\big) = \mathrm{Tr}_{s+1,\dots,N}\big(\partial_t f_N\big) = \mathrm{Tr}_{s+1,\dots,N}\big(\Omega f_N\big), \tag{6.1}$$

where we recall that Ω is the linear operator introduced in (1.10). Hence, due to (6.1), linearity of the trace map, and the definition of Ω, it follows that

$$\partial_t f_N^{(s)} = \sum_{K=1}^{M} \beta_K \frac{N}{K!\binom{N}{K}} \sum_{i_1\cdots i_K} \mathrm{Tr}_{s+1,\dots,N}\big(\Omega_{i_1\cdots i_K} f_N\big), \tag{6.2}$$

where for each $K = 1, \dots, M$ and $(i_1, \dots, i_K) \in \mathcal{I}(K)$, defined in (1.7) we have introduced the operator

$$\Omega_{i_1\cdots i_K} f = \int_{\mathbb{S}_K} \big(f \circ T^{\omega}_{i_1\cdots i_K} - f \big)\, \mathrm{d}b_K(\omega), \quad f \in L^1(\mathbb{R}^{dN}). \tag{6.3}$$

Thus, it remains to calculate the quantity $\mathrm{Tr}_{s+1,\dots,N}(\Omega_{i_1\cdots i_K} f)$ for arbitrary $(i_1, \dots, i_K) \in \mathcal{I}(K)$ and $f \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$.

The first step in this direction is exploiting the symmetric condition given in (1.3) in **(H3)**. This is the content of the following lemma. Recall that S_K stands for the group of permutations of K elements.

Lemma 6.1. *For all $K = 1, \dots, M$, $(i_1, \dots, i_K) \in \mathcal{I}(K)$, and $\gamma \in S_K$, it holds that*

$$\Omega_{i_{\gamma(1)}\cdots i_{\gamma(K)}} = \Omega_{i_1\cdots i_K}. \tag{6.4}$$

Proof. We divide the proof into two steps. In the first one, we assume that the collection of indices is a permutation of the first K indices: $\{1, \dots, K\}$. In the second step, we show how the general case follows from the particular one.

Step one. Let $\gamma \in S_K$ be any permutation of the elements $\{1, \dots, K\}$, and denote by $\Gamma = \gamma \times id_{N-K}$ its natural extension to S_N. Let $f \in L^1_{\mathrm{sym}}(\mathbb{R}^{dN}) \cap C(\mathbb{R}^{dN})$, and denote by $\Omega^+ = \Omega + id$ the *gain term* of (6.3). Then we calculate that for all $V \in \mathbb{R}^{dN}$

$$[\Omega^+_{1\cdots K} f](V) = \int_{\mathbb{S}_K} f\left[T^\omega_{1\cdots K} V\right] db_K(\omega) \tag{6.5}$$
$$= \int_{\mathbb{S}_K} f\left[T^\omega_K(v_1,\ldots,v_K); v_{K+1},\ldots,v_N\right] db_K(\omega) \tag{6.6}$$
$$= \int_{\mathbb{S}_K} f\left[\left(\gamma^{-1}\circ T^\omega_K \circ \gamma\right)(v_1,\ldots,v_K); v_{K+1},\ldots,v_N\right] db_K(\omega) \tag{6.7}$$
$$= \int_{\mathbb{S}_K} f\left[\left(\Gamma^{-1}\circ (T^\omega_K \times id_{\mathbb{R}^{d(N-K)}})\circ \Gamma\right) V\right] db_K(\omega) \tag{6.8}$$
$$= \int_{\mathbb{S}_K} f\left[T^\omega_{\gamma(1)\cdots\gamma(K)} V\right] db_K(\omega) \tag{6.9}$$
$$= [\Omega^+_{\gamma(1)\cdots\gamma(K)} f](V), \tag{6.10}$$

where we have used **(H3)** to obtain (6.7). Since $L^1_{\text{sym}} \cap C$ is a dense subspace of L^1_{sym}, this finishes the proof of the first step.

Step two. Now let $(i_1,\ldots,i_K) \in \mathcal{I}(K)$ be arbitrary and consider $\gamma \in S_K$ and $\Gamma \in S_N$ as in step one. First, we make a general observation: for all $\sigma \in S_N$, $f \in L^1_{\text{sym}}(\mathbb{R}^{dN})$, and $V \in \mathbb{R}^{dN}$, the following identity holds for the associated gain term:

$$\begin{aligned}[\Omega^+_{\sigma(1)\cdots\sigma(K)} f](V) &= \int_{\mathbb{S}_K} f\left[\left(\sigma\circ (T^\omega_K \times id_{\mathbb{R}^{d(N-K)}})\circ \sigma^{-1}\right) V\right] db_K(\omega) \\ &= [\Omega^+_{1\cdots K}(f\circ\sigma)](\sigma^{-1}V).\end{aligned} \tag{6.11}$$

Consequently, the same identity holds for the full operator as well. Now we choose σ such that $\sigma(1) = i_1, \ldots, \sigma(K) = i_K$. Then, step one and the general observation imply that

$$[\Omega_{i_1\cdots i_K} f](V) = [\Omega_{1\cdots K}(f\circ\sigma)](\sigma^{-1}V) \tag{6.12}$$
$$= [\Omega_{\gamma(1)\cdots\gamma(K)}(f\circ\sigma)](\sigma^{-1}V) \tag{6.13}$$
$$= [\Omega_{1\cdots K}(f\circ\sigma\circ\Gamma)]\left((\Gamma^{-1}\circ\sigma^{-1})V\right) \tag{6.14}$$
$$= [\Omega_{\sigma(\gamma(1))\cdots\sigma(\gamma(K))} f](V). \tag{6.15}$$

Since $\sigma(\gamma(\ell)) = i_{\gamma(\ell)}$ for all $\ell \in \{1,\ldots,K\}$, the proof is complete. ∎

We apply Lemma 6.1 in order to get a simplified expression of Ω. More precisely, we obtain that for all $K = 1,\ldots,M$ it holds that

$$\sum_{i_1\cdots i_K} \Omega_{i_1\cdots i_K} = \sum_{i_1<\cdots<i_K} \sum_{\mu\in S_K} \Omega_{i_{\mu(1)}\cdots i_{\mu(K)}} = K! \sum_{i_1<\cdots<i_K} \Omega_{i_1\cdots i_K}. \tag{6.16}$$

Consequently, we may plug this back in (6.2) to conclude that

$$\partial_t f^{(s)}_N = \sum_{K=1}^{M} \beta_K \frac{N}{\binom{N}{K}} \sum_{i_1<\cdots<i_K} \text{Tr}_{s+1,\ldots,N}\left(\Omega_{i_1\cdots i_K} f_N\right). \tag{6.17}$$

Thus, it suffices to calculate $\text{Tr}_{s+1,\ldots,N}(\Omega_{i_1\cdots i_K} f)$ only for ordered indices $i_1 < \cdots < i_K$ and symmetric functions $f \in L^1_{\text{sym}}(\mathbb{R}^{dN})$. The following family of operators is defined with that purpose.

DEFINITION 6. *Let $K = 1,\ldots,M$, let $n = 1,\ldots,K$, and denote $r \equiv K - n$. For all indices $1 \le i_1 < \cdots < i_n \le s$, we define the operator*

$$C^{s,K,n}_{i_1\cdots i_n} : L^1_{\text{sym}}(\mathbb{R}^{d(s+r)}) \to L^1_{\text{sym}}(\mathbb{R}^{ds})$$

as follows:

1. *For* $r=0$, *we set* $C^{s,K,n}_{i_1\cdots i_n} := \Omega_{i_1\cdots i_K}$.
2. *For* $r\geq 1$ *and* $s+r\leq N$, *we set*

$$C^{s,K,n}_{i_1\cdots i_n} f^{(s+r)}(V_s) := \int_{\mathbb{S}_K\times\mathbb{R}^{dr}} \big(f^{(s+r)}(V^{*i_1\cdots *i_n}_{s+r}) - f^{(s+r)}(V_{s+r})\big)\mathrm{d}b_K(\omega)dv_{s+1}\cdots dv_{s+r},$$

where $V_s\in\mathbb{R}^{ds}$, $V_{s+r}\equiv(V_s,v_{s+1},\ldots,v_{s+r})\in\mathbb{R}^{d(s+r)}$, *and*

$$V^{*i_1\cdots *i_n}_{s+r} := (v_1,\ldots,v^*_{i_1},\ldots,v^*_{i_n},\ldots,v_s\,;\,v^*_{s+1},\ldots,v^*_{s+r})\in\mathbb{R}^{d(s+r)} \tag{6.18}$$

with $(v^*_{i_1},\ldots,v^*_{i_n},v^*_{s+1},\ldots,v^*_{s+r}) = T^\omega_K(v_{i_1},\ldots,v_{i_n},v_{s+1},\ldots,v_{s+r})\in\mathbb{R}^{dK}$.

3. *For* $r\geq 1$ *and* $s+r>N$, *we set* $C^{s,K,n}_{i_1\cdots i_n}\equiv 0$.

Remarks 6.1. A few comments are in order.

(i) Let us briefly try to motivate the involved notation. As we shall see, one of the main results of this section is Lemma 6.5, and its proof relies on a careful classification of small and large terms. In particular, the index n introduced in Definition 6 helps with this classification. Namely, the operators with $n=1$ give rise to the leading-order contributions. The $n\geq 2$ terms give rise to remainder terms. See also Remark 6.3.

(ii) It can be helpful to keep in mind that s is the order of the marginal and r is the number of interacting particles that get traced over by the operator $C^{s,K,n}_{i_1\cdots i_n}$.

(iii) $C^{s,K,n}_{i_1\cdots i_n}$ is bounded with operator norm $\|C^{s,K,n}_{i_1\cdots i_n}\|\leq 2$.

The following lemma is the main result concerning the operators just introduced.

LEMMA 6.2. *Let* $K=1,\ldots,M$, *let* $n=1,\ldots,K$, *and let* $r\equiv K-n$. *Assume that* $s+r\leq N$, *and consider* K *ordered indices such that*

$$1\leq i_1<\cdots<i_n\leq s<i_{n+1}<\cdots<i_{n+r}=i_K\leq N. \tag{6.19}$$

Then, for all $f\in L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$, *the following identity holds:*

$$Tr_{s+1,\ldots,N}\big(\Omega_{i_1\cdots i_K}f\big) = C^{s,K,n}_{i_1\cdots i_n}[Tr_{s+r+1,\ldots,N}f]\,. \tag{6.20}$$

Remark 6.1. The main consequence of the previous result is that the left-hand side of (6.20) is independent of the last r indices $(i_{n+1},\ldots,i_{n+r})$.

Proof. Since $s+r\leq N$, there are two cases.
(i) Let $r=0$. Then $i_1<\cdots<i_K\leq s$. In particular, all of the particles that are being traced out are not interacting. Consequently, it is easy to show that

$$\mathrm{Tr}_{s+1,\ldots,N}\big(\Omega_{i_1\cdots i_K}f\big) = \Omega_{i_1\cdots i_K}[\mathrm{Tr}_{s+1,\ldots,N}f] = C^{s,K,n}_{i_1\cdots i_n}[\mathrm{Tr}_{s+1,\ldots,N}f]\ . \tag{6.21}$$

(ii) Let $r\geq 1$. Let $f\in L^1_{\mathrm{sym}}(\mathbb{R}^{dN})$, and fix $V_s\in\mathbb{R}^{ds}$. Let μ be any permutation of the elements $\{s+1,\ldots,N\}$. Then we may implement the change of variables $(v_{s+1},\ldots,v_N)\mapsto(v_{\mu^{-1}(s+1)},\ldots,v_{\mu^{-1}(N)})$ in the following expression:

$$\begin{aligned}
&\int_{\mathbb{R}^{d(N-s)}} f\big[T^\omega_{i_1\cdots i_K}(V_s;v_{s+1},\ldots,v_N)\big]\mathrm{d}v_{s+1},\ldots,\mathrm{d}v_N\\
&\quad=\int_{\mathbb{R}^{d(N-s)}} f\big[T^\omega_{i_1\cdots i_K}(V_s;v_{\mu^{-1}(s+1)},\ldots,v_{\mu^{-1}(N)})\big]\mathrm{d}v_{s+1},\ldots,\mathrm{d}v_N\\
&\quad=\int_{\mathbb{R}^{d(N-s)}} f\big[\big(T^\omega_{i_1\cdots i_K}\circ(id_s\times\mu^{-1})\big)(V_s;v_{s+1},\ldots,v_N)\big]\mathrm{d}v_{s+1},\ldots,\mathrm{d}v_N,
\end{aligned} \tag{6.22}$$

where we recall that we identify $id_s \times \mu^{-1}$ with its group action over $\mathbb{R}^{dN}$, i.e., we write

$$(id_s \times \mu^{-1})(V_s; v_{s+1}, \dots, v_N) = (V_s; v_{\mu^{-1}(s+1)}, \dots, v_{\mu^{-1}(N)}).$$

Next, since $f \in L^1_{\text{sym}}$, there holds that $f = f \circ \bar{\mu}$, where we denote $\bar{\mu} \equiv id_s \times \mu \in S_N$. Therefore, denoting $V \equiv (V_s, v_{s+1}, \dots, v_N)$, we obtain, thanks to (6.22) and permutational symmetry, that

$$\begin{aligned}\int_{\mathbb{R}^{d(N-s)}} f\big[T^{\omega}_{i_1\cdots i_K} V\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N &= \int_{\mathbb{R}^{d(N-s)}} f\big[(\bar{\mu} \circ T^{\omega}_{i_1\cdots i_K} \circ \bar{\mu}^{-1})\, V\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N \\ &= \int_{\mathbb{R}^{d(N-s)}} f\big[T^{\omega}_{\bar{\mu}(i_1)\cdots\bar{\mu}(i_K)} V\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N,\end{aligned} \tag{6.23}$$

where the last line follows from the definition of $T^{\omega}_{i_1\dots i_K}$ (see (1.8)) upon conjugation with $\bar{\mu}$. Since $1 \le i_1 < \cdots < i_n \le s$, we must have $\bar{\mu}(i_\ell) = i_\ell$ for $1 \le \ell \le n$. Further, since $s+1 \le i_{n+1} < \cdots < i_{n+r} \le N$, we may choose μ such that $\mu(i_{n+1}) = s+1, \dots, \mu(i_{n+r}) = \mu(i_K) = s + r$. Consequently, we find that

$$\int_{\mathbb{R}^{d(N-s)}} f\big[T^{\omega}_{i_1\cdots i_K} V\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N = \int_{\mathbb{R}^{d(N-s)}} f\big[T^{\omega}_{i_1\dots i_n, s+1\cdots s+r} V\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N. \tag{6.24}$$

Next, using the notation introduced in Definition 6, we are able to write

$$T^{\omega}_{i_1\dots i_n, s+1\cdots s+r} V = (V^{*i_1\cdots *i_n}_{s+r}, v_{s+r+1}, \dots, v_N)\ . \tag{6.25}$$

Hence, we may use Fubini's theorem over the space $\mathbb{R}^{d(N-s)} = \mathbb{R}^{dr} \times \mathbb{R}^{d(N-s-r)}$ to find that

$$\begin{aligned}&\int_{\mathbb{R}^{d(N-s)}} f\big[T^{\omega}_{i_1\cdots i_K} V\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N \\ &= \int_{\mathbb{R}^{d(N-s)}} f\big[V^{*i_1\cdots *i_n}_{s+r}, v_{s+r+1}, \dots, v_N\big] \mathrm{d}v_{s+1}, \dots, \mathrm{d}v_N \\ &= \int_{\mathbb{R}^{dr}} \Big(\int_{\mathbb{R}^{d(N-r-s)}} f\big[V^{*i_1\cdots *i_n}_{s+r}, v_{s+r+1}, \dots, v_N\big]\, \mathrm{d}v_{s+r+1} \cdots \mathrm{d}v_N \Big) \mathrm{d}v_{s+1} \cdots \mathrm{d}v_{s+r} \\ &= \int_{\mathbb{R}^{dr}} f^{(s+r)}[V^{*i_1\cdots *i_n}_{s+r}]\, \mathrm{d}v_{s+1} \cdots \mathrm{d}v_{s+r},\end{aligned} \tag{6.26}$$

where, in order to obtain the last line, we have used the definition of the marginals introduced in section 2. Similarly, one can prove that $\int_{\mathbb{R}^{d(N-s)}} f[V] \mathrm{d}v_{s+1} \dots \mathrm{d}v_N = \int_{\mathbb{R}^{dr}} f^{(s+r)}[V_{s+r}] \mathrm{d}v_{s+1} \cdots \mathrm{d}v_{s+r}$. We subtract these two identities and integrate against $\mathrm{d}b_K(\omega)$ to prove our claim. ∎

Definition 7. *For $N \ge M$, $1 \le s \le N$, and $0 \le k \le m$, we define the linear operator*

$$\mathcal{C}^N_{s,s+k} : L^1_{\text{sym}}(\mathbb{R}^{d(s+k)}) \to L^1_{\text{sym}}(\mathbb{R}^{ds})$$

according to the formula

$$\mathcal{C}^N_{s,s+k} := \sum_{n=1}^{M-k} \beta_{k+n} \frac{N}{\binom{N}{k+n}} \binom{N-s}{k} \sum_{1 \le i_1 < \cdots < i_n \le s} C^{s,k+n,n}_{i_1,\dots,i_n}\ . \tag{6.27}$$

Remark 6.2. It is straightforward to verify that, for each $N \in \mathbb{N}$, there is only finitely many operators that are nonzero. In particular, $\mathcal{C}^N_{s,s+k} = 0$ for any $s \in \mathbb{N}$ satisfying $s + k > N$.

We are now ready to record the BBGKY hierarchy.

LEMMA 6.3. *For all $N \geq M$ and $1 \leq s \leq N$, let f_N denote the solution of the master equation* (1.10) *and let $(f_N^{(s)})_{s\in\mathbb{N}}$ be its sequence of marginals, defined in* (2.21). *Then it holds that*

$$\partial_t f_N^{(s)} = \sum_{k=0}^{m} \mathcal{C}^N_{s,s+k} f_N^{(s+k)}. \tag{6.28}$$

Proof. First, following the same argument of the proof of Lemma 6.2, we may verify that for $s < i_1 < \cdots < i_K$ it holds that

$$\operatorname{Tr}_{s+1,\ldots,N}\big(\Omega_{i_1\cdots i_K} f\big) = 0. \tag{6.29}$$

Next, we use the following decomposition of the set of ordered indices:

$$\sum_{i_1<\cdots<i_K} \operatorname{Tr}_{s+1,\ldots,N}\big(\Omega_{i_1\cdots i_K} f\big) = \sum_{n=1}^{K} \sum_{\substack{i_1<\cdots<i_K \\ i_n\leq s<i_{n+1}}} \operatorname{Tr}_{s+1,\ldots,N}\big(\Omega_{i_1\cdots i_K} f\big), \tag{6.30}$$

where for notational convenience we denote $i_{K+1} = N + 1$. In other words, n counts the number of the indices $\{i_\ell\}$ that are less than or equal to s. In addition, we note that

$$s < i_{n+1} < \cdots < i_K \leq N \quad \Longrightarrow \quad N \geq s + (K - n). \tag{6.31}$$

We implement (6.31) by means of a characteristic function $\mathbb{1}_{K,n} \equiv \mathbb{1}(N \geq s + K - n)$. Thus, we may write, thanks to (6.17) and Lemma 6.2,

$$\begin{aligned}
\partial_t f_N^{(s)} &= \sum_{K=1}^{M} \beta_K \frac{N}{\binom{N}{K}} \sum_{n=1}^{K} \mathbb{1}_{K,n} \sum_{\substack{i_1<\cdots<i_K \\ i_n\leq s<i_{n+1}}} \operatorname{Tr}_{s+1,\ldots,N}\big(\Omega_{i_1\cdots i_K} f_N\big) \\
&= \sum_{K=1}^{M} \beta_K \frac{N}{\binom{N}{K}} \sum_{n=1}^{K} \mathbb{1}_{K,n} \sum_{\substack{i_1<\cdots<i_K \\ i_n\leq s<i_{n+1}}} C^{s,K,n}_{i_1\cdots i_n} f_N^{(s+K-n)}.
\end{aligned} \tag{6.32}$$

Note that $C^{s,K,n}_{i_1\cdots i_n} f_N^{(s+K-n)}$ does not depend on the indices $i_{n+1} < \cdots < i_{n+r}$, so these can be summed out. We find that

$$\begin{aligned}
\partial_t f_N^{(s)} &= \sum_{K=1}^{M} \beta_K \frac{N}{\binom{N}{K}} \sum_{n=1}^{K} \mathbb{1}_{K,n} \sum_{1\leq i_1<\cdots<i_n\leq s} \left(\sum_{s+1\leq i_{n+1}<\cdots i_K\leq N} 1 \right) C^{s,K,n}_{i_1\cdots i_n} f_N^{(s+K-n)} \\
&= \sum_{K=1}^{M} \beta_K \frac{N}{\binom{N}{K}} \sum_{n=1}^{K} \mathbb{1}_{K,n} \sum_{1\leq i_1<\cdots<i_n\leq s} \binom{N-s}{K-n} C^{s,K,n}_{i_1\cdots i_n} f_N^{(s+K-n)}.
\end{aligned} \tag{6.33}$$

Further, note that $\mathbb{1}_{K,n} C^{s,K,n}_{i_1\cdots i_n} = C^{s,K,n}_{i_1\cdots i_n}$. Finally, we make the substitution $r = K - n$ to obtain

$$\partial_t f_N^{(s)} = \sum_{r=0}^{M-1} \sum_{n=1}^{M-r} \beta_{r+n} \frac{N}{\binom{N}{r+n}} \binom{N-s}{r} \sum_{1\le i_1 < \cdots < i_n \le s} C^{s,r+n,n}_{i_1\cdots i_n} f_N^{(s+r)}$$

$$= \sum_{r=0}^{m} \mathcal{C}^N_{s,s+r} f_N^{(s+r)}, \tag{6.34}$$

where on the last line we recall that $M-1=m$. This finishes the proof. ∎

The next result shows that the operators that drive the BBGKY hierarchy fit the abstract framework introduced in section 2.

LEMMA 6.4. *Assume $M/N \le \varepsilon \in (0,1)$. Then the operators $(\mathcal{C}^N_{s,s+k})_{k=0}^m$ satisfy Condition 1 with constants $(R_k)_{k=0}^m$ given by*

$$R_k = 2 \sum_{\ell=k+1}^{M} \frac{\beta_\ell}{(1-\varepsilon)^\ell} \binom{\ell}{k}. \tag{6.35}$$

In the upcoming proof, we will make use of the following two inequalities:

$$(1-k/n)^k \frac{n^k}{k!} \le \binom{n}{k} \le \frac{n^k}{k!} \qquad \forall n \in \mathbb{N},\ \forall k \le n, \tag{6.36}$$

which can be easily derived by noting that

$$(1-k/n)^k n^k = (n-k)^k \le n(n-1)\cdots(n-(k-1)) \le n^k. \tag{6.37}$$

Proof. We assume without loss of generality that $s \le N$, for otherwise $\mathcal{C}^N_{s,s+k} = 0$ for any $k \ge 0$. First, recalling that $\|C^{s,K,n}_{i_1\cdots i_n}\| \le 2$ we find that

$$\|\mathcal{C}^N_{s,s+k}\| = \left\| \sum_{n=1}^{M-k} \beta_{k+n} \frac{N}{\binom{N}{k+n}} \binom{N-s}{k} \sum_{1\le i_1<\cdots<i_n\le s} C^{s,k+n,n}_{i_1,\dots,i_n} \right\| \tag{6.38}$$

$$\le 2 \sum_{n=1}^{M-k} \beta_{k+n} \frac{N}{\binom{N}{k+n}} \binom{N-s}{k} \binom{s}{n}, \tag{6.39}$$

where we have used the fact that $\sum_{1\le i_1<\cdots<i_n\le s} = \binom{s}{n}$. Next, we use (6.36) to estimate

$$N \binom{N-s}{k} \binom{s}{n} \le N \binom{N}{k} \binom{s}{n} \le \frac{N^{k+1}}{k!} \frac{s^n}{n!}\ . \tag{6.40}$$

Similarly, for the denominator we find that

$$\binom{N}{n+k} \ge \Big(1-(n+k)/N\Big)^{n+k} \frac{N^{n+k}}{(n+k)!} \ge (1-\varepsilon)^{n+k} \frac{N^{n+k}}{(n+k)!}, \tag{6.41}$$

where, for the second inequality, we used the fact that $(n+k)/N \le M/N \le \varepsilon$. We put together (6.39), (6.40), and (6.41) to find that

$$\|\mathcal{C}^N_{s,s+k}\| \ \le\ 2 \sum_{n=1}^{M-k} \beta_{k+n} (1-\varepsilon)^{-(n+k)} \binom{n+k}{k} \frac{s^n}{N^{n-1}} \ \le\ R_k\, s, \tag{6.42}$$

where, in the second inequality, we have used the upper bound $s^n N^{-(n-1)} \le s$, followed by a change of variables $\ell = n+k$. This finishes the proof. ∎

6.2. Convergence of operators. For $s \in \mathbb{N}$ and $0 \le k \le M$, we introduce the operator

$$\mathcal{C}^{\infty}_{s,s+k} : L^1_{\mathrm{sym}}(\mathbb{R}^{d(s+k)}) \to L^1_{\mathrm{sym}}(\mathbb{R}^{ds})$$

given by

$$\begin{aligned}&(\mathcal{C}^{\infty}_{s,s+k} f^{(s+k)})(V_s)\\ &\quad := \beta_K K \sum_{i=1}^{s} \int_{\mathbb{S}_K \times \mathbb{R}^{dk}} \Big(f^{(s+k)}(V^{*i}_{s+k}) - f^{(s+k)}(V_{s+k}) \Big) \mathrm{d}b_K(\omega) \mathrm{d}v_{s+1} \cdots v_{s+k}\end{aligned}$$

for $k \ge 1$ and with the obvious modification for $k = 0$. Here V^*_{s+k} is as in Definition 6.

Our following result establishes convergence of operators, which in turn allow us to apply Theorem 2.1. In order to state it, we introduce on $X = \oplus_{s\in\mathbb{N}} L^1_{\mathrm{sym}}(\mathbb{R}^{ds})$ the linear operators

$$(\mathcal{C}^N F)^{(s)} := \sum_{k=0}^{m} \mathcal{C}^N_{s,s+k} f^{(s+k)}, \qquad F = (f^{(s)})_{s\in\mathbb{N}}, \tag{6.43}$$

where $\mathcal{C}^N_{s,s+k}$ was defined in Definition 7, and

$$(\mathcal{C}^{\infty} F)^{(s)} := \sum_{k=0}^{m} \mathcal{C}^{\infty}_{s,s+k} f^{(s+k)}, \qquad F = (f^{(s)})_{s\in\mathbb{N}}. \tag{6.44}$$

Lemma 6.5. *Let $\mathcal{C}^N$ be as in* (6.43) *and $\mathcal{C}^{\infty}$ be as in* (6.44)*, respectively. Then, $\mathcal{C}^N$ converges to $\mathcal{C}^{\infty}$ in the sense of Definition* 4.

Remark 6.3 (heuristics). Let us briefly informally explain the motivation for the proof of Lemma 6.5. To this end, we consider the following simplified system of equations as a prototypical example:

$$\partial_t f^{(s)}_N = \frac{1}{N} \sum_{1 \le i_1 \le i_2 \le N} E^{s,s+1}_{i_1,i_2} f^{(s+1)}_N \tag{6.45}$$

for some linear operators $E^{s,s+1}_{i,j}$. Here $N \ge 1$ is large, and $1 \le s \le N$ is the "order" of the marginals, and $1 \le i_1, i_2 \le N$ label the indices of the interacting particles. One can then split the sums as follows:

$$\frac{1}{N} \sum_{1 \le i_1 \le i_2 \le N} E^{s,s+1}_{i_1,i_2} = \frac{1}{N} \sum_{1 \le i_1 \le s \le i_2 \le N} E^{s,s+1}_{i_1,i_2} + \frac{1}{N} \sum_{1 \le i_1 \le i_2 \le s} E^{s,s+1}_{i,j} \tag{6.46}$$

if we also assume that $E^{s,s+1}_{i_1,i_2} = 0$ for $i_1, i_2 > s$. A counting argument shows that the first term is $O(1)$ and contributes to leading order. The second term is $O(1/N)$ and vanishes in the limit. The upcoming proof separates the indices $i_1 \le \cdots \le i_K$ in the same spirit, according to the index $n \ge 1$, which labels the number of indices below s.

Proof. Let us fix $k \in \{0, \dots, M-1\}$. First, we decompose the BBGKY operator into a *leading-order term* and a *remainder term:*

$$\mathcal{C}^N_{s,s+k} = \tilde{\mathcal{C}}^N_{s,s+k} + \mathcal{R}^N_{s,s+k}. \tag{6.47}$$

This decomposition follows from (6.27). The leading-order term corresponds to the $n=1$ contribution, whereas the remainder term corresponds to the $n\geq 2$ contribution. Explicitly, we have

$$\widetilde{\mathcal{C}}^N_{s,s+k} := \beta_{k+1}\frac{N}{\binom{N}{k+1}}\binom{N-s}{k}\sum_{1\leq i\leq s} C_i^{s,k+1,1} \tag{6.48}$$

and

$$\mathcal{R}^N_{s,s+k} := \sum_{n=2}^{M-k}\beta_{k+n}\frac{N}{\binom{N}{k+n}}\binom{N-s}{k}\sum_{1\leq i_1<\cdots<i_n\leq s} C^{s,k+n,n}_{i_1,\dots,i_n}. \tag{6.49}$$

The following is enough to prove our claim. Let $F_N=(f_N^{(s)})_{s\in\mathbb{N}}\in X$ converge weakly to $F=(f^{(s)})_{s\in\mathbb{N}}\in X$. Then, for all $s\in\mathbb{N}$ and $\varphi_s\in L^\infty(\mathbb{R}^{ds})$, we have the following:

1. There holds that

$$\lim_{N\to\infty}\left\langle \widetilde{\mathcal{C}}^N_{s,s+k}f_N^{(s+k)},\varphi_s\right\rangle = \left\langle \mathcal{C}^\infty_{s,s+k}f^{(s+k)},\varphi_s\right\rangle.$$

2. There holds that

$$\lim_{N\to\infty}\left\langle \mathcal{R}^N_{s,s+k}f_N^{(s+r)},\varphi_s\right\rangle = 0.$$

We shall assume for simplicity that $k\geq 1$, the case $k=0$ being analogous.

Proof of item 1. Let us denote by $\mathcal{D}_{s+k,s}=(\mathcal{C}^\infty_{s,s+k})^*: L^\infty(\mathbb{R}^{ds})\to L^\infty(\mathbb{R}^{d(s+k)})$ the Banach space adjoints of the limiting collisional operators. In particular, they admit the representation

$$(\mathcal{D}_{s+k,s}\varphi_s)(V_{s+k}) = \beta_{k+1}(k+1)\sum_{i=1}^s\int_{\mathbb{S}_K}\Big((\varphi_s\otimes\mathbb{1}_k)(V^{*i}_{s+k})-(\varphi_s\otimes\mathbb{1}_k)(V_{s+k})\Big)\mathrm{d}b_K(\omega),$$

where $\mathbb{1}_k$ is the dk-dimensional identity. A straightforward calculation based on a change of variables shows that

$$\langle\widetilde{\mathcal{C}}^N_{s,s+k}f_N^{(s+k)},\varphi\rangle = \frac{1}{k+1}\frac{N\binom{N-s}{k}}{\binom{N}{k+1}}\left\langle f_N^{(s+k)},\mathcal{D}_{s+k,s}\varphi_s\right\rangle. \tag{6.50}$$

Since $\mathcal{D}_{s+k,s}\varphi_s\in L^\infty(\mathbb{R}^{d(s+k)})$, we use weak convergence of the marginals to calculate that

$$\begin{aligned}\lim_{N\to\infty}\langle\widetilde{\mathcal{C}}^N_{s,s+k}f_N^{(s+k)},\varphi\rangle &= \lim_{N\to\infty}\left(\frac{1}{k+1}\frac{N\binom{N-s}{k}}{\binom{N}{k+1}}\right)\lim_{N\to\infty}\langle f_N^{(s+k)},\mathcal{D}_{s+k,s}\varphi_s\rangle\\ &=\langle f^{(s+k)},\mathcal{D}_{s+k,s}\varphi_s\rangle\\ &=\langle\mathcal{C}^\infty_{s,s+k}f^{(s+k)},\varphi_s\rangle. \qquad (6.51)\end{aligned}$$

This finishes the proof of item 1.

Proof of item 2. First, we establish a norm estimate for the remainder term. The same analysis done in Lemma 6.4 can be carried out for the remainder term to find that for $N\geq\varepsilon^{-1}M$

$$\|\mathcal{R}^N_{s,s+k}\|\leq 2\sum_{n=2}^{M-k}\frac{\beta_{n+k}}{(1-\varepsilon)^{n+k}}\binom{n+k}{k}\frac{s^n}{N^{n-1}}\ . \tag{6.52}$$

For $s \le N$ and $n \ge 2$, we can now use the alternative upper bound $s^n = s^2 s^{n-2} \le s^2 N^{n-2}$ to find that the following estimate holds:

$$\|\mathcal{R}^N_{s,s+k}\| \le C_k \frac{s^2}{N}, \tag{6.53}$$

where $C_k = \sum_{\ell=k+2}^{M} (1-\varepsilon)^{-\ell} \beta_\ell \binom{\ell}{k}$. Next, fix $s \in \mathbb{N}$ and note that, thanks to weak convergence and the uniform boundedness principle, the quantity $K_s = \sup_{N\in\mathbb{N}} \|f_N^{(s)}\|_{L^1_{\mathrm{sym}}(\mathbb{R}^{ds})}$ is finite. Thus, we find that the following estimate holds:

$$|\langle \mathcal{R}^N_{s,s+k} f^N_{s,s+k}, \varphi_s \rangle| \le K_s \|\mathcal{R}^N_{s,s+k}\| \, \|\varphi_s\|_{L^\infty} \le \frac{K_s C_k s^2}{N} \|\varphi_s\|_{L^\infty}, \tag{6.54}$$

from which our claim follows after taking the $N \to \infty$ limit. ◻

Proof of Theorem 2.2. Lemma 6.4 implies that the operator $\mathcal{C}^N$ satisfies Condition 1. Similar arguments show that $\mathcal{C}^\infty$ satisfies Condition 2. Further, Lemma 6.5 shows that $\mathcal{C}^N$ converges to $\mathcal{C}^\infty$ in the sense of Definition 4. In order to prove Theorem 2.2, it suffices to apply Theorem 2.1 to any mild solutions of the BBGKY and Boltzmann hierarchies, respectively. ◻

Proof of Theorem 2.3. Let f_N be the solution of the master equation (1.10), $(f_N^{(s)})_{s\in\mathbb{N}}$ its sequence of marginals (2.21), and $f_0 \in L^1(\mathbb{R}^d)$ the initial datum for which $f_N^{(s)}(0)$ converges pointwise weakly to $f_0^{\otimes s}$. We apply Theorem 2.2 to conclude that $(f_N^{(s)})_{s\in\mathbb{N}}$ converges in observables to $F = (f^{(s)})_{s\in\mathbb{N}}$—the solution of the Boltzmann hierarchy (2.27) with initial data $F_0 = (f_0^{\otimes s})_{s\in\mathbb{N}}$—over $[0,T]$.

Finally, let $f(t,v)$ be the solution of the generalized Boltzmann equation (2.28) with initial data f_0. A straightforward calculation shows that $(f^{\otimes s})_{s\in\mathbb{N}}$ is a mild solution of the Boltzmann hierarchy (2.27). Because of Proposition 2.2, the Boltzmann hierarchy is well-posed. Uniqueness then implies that $f^{(s)} = f^{\otimes s}$ for all $s \in \mathbb{N}$. Consequently, for all $\varphi_s \in L^\infty(\mathbb{R}^{ds})$ it holds that

$$\langle f_N^{(s)}(t), \varphi_s \rangle \longrightarrow \langle f(t)^{\otimes s}, \varphi_s \rangle \quad \text{as} \quad N \to \infty$$

uniformly in $t \in [0,T]$. Since T is independent of the initial conditions and thanks to the global a priori bounds

$$\sup_{N\in\mathbb{N}} \sup_{s\le N} \|f_N^{(s)}(t,\cdot)\|_{L^1} = 1 \qquad \forall t \ge 0, \tag{6.55}$$

one may repeat the above argument to prove convergence for arbitrarily large $t \ge 0$. This finishes the proof. ◻

7. Well-posedness.

7.1. The hierarchies. In this subsection, we address the question of well-posedness of the finite and infinite hierarchies, respectively. We only give a proof of Proposition 2.2, the other one being completely analogous.

Proof of Proposition 2.2. For simplicity, let us denote $\mathscr{C} \equiv \mathscr{C}^\infty$. Let $\boldsymbol{F} = (\boldsymbol{f}^{(s)})_{s\in\mathbb{N}} \in \boldsymbol{X}_{\boldsymbol{\mu}}$. Then we obtain thanks to Condition 2 the following estimate:

$$
\begin{aligned}
\left\| \left(\int_0^t \mathscr{C}\big[\boldsymbol{F}(\tau)\big]\, \mathrm{d}\tau \right)^{(s)} \right\|_{X^{(s)}} &\leq \int_0^t \|\Big(\mathscr{C}\big[\boldsymbol{F}(\tau)\big]\Big)^{(s)}\|_{X^{(s)}}\, \mathrm{d}\tau \\
&\leq \int_0^t \sum_{k=0}^{m} s\rho_k \|\boldsymbol{f}^{(s+k)}(\tau)\|_{X^{(s)}}\, \mathrm{d}\tau \\
&= \int_0^t \sum_{k=0}^{m} s\rho_k e^{-\boldsymbol{\mu}(\tau)(s+k)} e^{\boldsymbol{\mu}(\tau)(s+k)} \|\boldsymbol{f}^{(s+k)}(\tau)\|_{X^{(s)}}\, \mathrm{d}\tau \\
&\leq \int_0^t \sum_{k=0}^{m} s\rho_k e^{-\boldsymbol{\mu}(\tau)(s+k)}\, \mathrm{d}\tau\, \|\boldsymbol{F}\|_{\boldsymbol{\mu}} \\
&\leq T \sum_{k=0}^{m} \frac{s}{s+k} \rho_k e^{-\boldsymbol{\mu}(t)(s+k)}\, \|\boldsymbol{F}\|_{\boldsymbol{\mu}} \\
&\leq \left(T \sum_{k=0}^{m} \rho_k e^k \right) e^{-\boldsymbol{\mu}(t)s} \|\boldsymbol{F}\|_{\boldsymbol{\mu}} = \theta e^{-\boldsymbol{\mu}(t)s} \|\boldsymbol{F}\|_{\boldsymbol{\mu}}, \qquad (7.1)
\end{aligned}
$$

where we have defined $\theta := T\sum_{k=0}^{m} \rho_k e^k \in (0,1)$.

On $\boldsymbol{X}_{\boldsymbol{\mu}}$ we introduce the map $\boldsymbol{F} \mapsto F_0 + \int_0^t \mathscr{C}\boldsymbol{F}(\tau)\mathrm{d}\tau =: \mathcal{M}[\boldsymbol{F}]$. Linearity of $\mathscr{C}$ and the estimate contained in (7.1) imply that $\|\mathcal{M}[\boldsymbol{F}] - \mathcal{M}[\boldsymbol{G}]\|_{\boldsymbol{\mu}} \leq \theta \|\boldsymbol{F} - \boldsymbol{G}\|_{\boldsymbol{\mu}}$ for all $\boldsymbol{F}, \boldsymbol{G} \in \boldsymbol{X}_{\boldsymbol{\mu}}$. Therefore, $\mathcal{M}$ is a contraction. Let $r = (1-\theta)^{-1}\theta \in (0,\infty)$, and define $R := r\|F_0\|_{X_0}$. Then estimate (7.1) and the triangle inequality show that

$$
\|\mathcal{M}[\boldsymbol{F}] - F_0\|_{\boldsymbol{\mu}} \leq \theta \|\boldsymbol{F}\|_{\boldsymbol{\mu}} \leq \theta \|\boldsymbol{F} - F_0\|_{\boldsymbol{\mu}} + \theta \|F_0\|_{\boldsymbol{\mu}} \leq R
$$

whenever $\|\boldsymbol{F} - F_0\|_{\boldsymbol{\mu}} \leq R$. Therefore, $\mathcal{M}$ maps the ball $B_R(F_0) \subset \boldsymbol{X}_{\boldsymbol{\mu}}$ or radius R around F_0 into itself. The conclusion of the theorem now follows from Banach's fixed point theorem. The continuity estimate also follows easily from our considerations. ∎

7.2. The Boltzmann equation. The main goal of this subsection is to prove Proposition 2.3. First, we prove the following two lemmas.

LEMMA 7.1 (local well-posedness). *For all $f_0 \in L^1(\mathbb{R}^d)$, there exists $0 < T_* = T_*(\|f_0\|_{L^1})$ such that there is a unique mild solution*

$$
f \in C\big([0,T_*], L^1(\mathbb{R}^d)\big) \cap C^1\big((0,T_*), L^1(\mathbb{R}^d)\big)
$$

to the Boltzmann equation (2.28) *with initial data f_0.*

Proof. The operators Q_K satisfy the following estimates: for $f, g \in L^1$, there holds that

$$
\|Q_K(f)\|_{L^1} \leq 2K\|f\|_{L^1}^K, \qquad (7.2)
$$

$$
\|Q_K(f) - Q_K(g)\|_{L^1} \leq 2K^2\big(\|f\|_{L^1}^{K-1} + \|g\|_{L^1}^{K-1}\big)\|f - g\|_{L^1}. \qquad (7.3)
$$

Thanks to these estimates, a proof based on a fixed-point argument on $C([0,T_*], L^1(\mathbb{R}^d))$ shows that there is a unique solution to the integral equation $f(t) = f_0 + \int_0^t \sum_{K=1}^{M} \beta_K Q_K[f(s),\ldots,f(s)]\mathrm{d}s$. We leave the details to the reader.

Finally, note that thanks to the estimate (7.3), it is easy to show that the map $t \mapsto Q_K[f(t),\ldots,f(t)] \in L^1(\mathbb{R}^d)$ is continuous. It then follows from the fundamental theorem of calculus that $f \in C^1\big((0,T_*), L^1\big)$ and (2.28) holds in the strong sense. ∎

LEMMA 7.2 (conservation of mass). *Let* $f \in C([0,T_*], L^1(\mathbb{R}^d)) \cap C^1\big((0,T_*), L^1(\mathbb{R}^d)\big)$ *be the continuous solution of the Boltzmann equation in mild form* (2.28). *Then*

$$\int_{\mathbb{R}^d} f(t,v)\mathrm{d}v = \int_{\mathbb{R}^d} f_0(v)\mathrm{d}v \qquad \forall t \in [0,T_*]. \tag{7.4}$$

Proof. Since (2.28) holds in the strong sense, we may calculate thanks to a change of variables that

$$\begin{aligned} \partial_t \int_{\mathbb{R}^d} f(t,v)\mathrm{d}v &= \int_{\mathbb{R}^d} \partial_t f(t,v)\mathrm{d}v \\ &= \sum_{K=1}^{M} \int_{\mathbb{R}^d} \beta_K Q_K[f(t),\dots,f(t)](v)\ \mathrm{d}v = 0 \qquad \forall t \in (0,T_*). \end{aligned} \tag{7.5}$$

This finishes the proof. ∎

Proof of Proposition 2.3. Let f be the solution to the Boltzmann equation (2.28), with initial data f_0 satisfying $\int_{\mathbb{R}^d} f_0(v)\mathrm{d}v = 1$ and $\|f_0\|_{L^1} \le 1$, whose existence is guaranteed by Lemma 7.1. Our goal will be to show that $\|f(t)\|_{L^1} \le 1$ for all $t \in (0,T_*)$, after possibly reducing T_* by a constant depending only on M and $\{\beta_K\}_{K=1}^M$. One may then patch the solutions obtained by Lemma 7.1 to obtain global well-posedness.

First, we note that thanks to conservation of mass, the collisional operators given in (1.6), when acting on f, may be written as

$$Q_K[f(t),\dots,f(t)] = Q_K^{(+)}[f(t),\dots,f(t)] - Kf(t), \qquad K=1,\dots,M, \tag{7.6}$$

where $Q_K^{(+)} : L^1(\mathbb{R}^d)^K \to L^1(\mathbb{R}^d)$ corresponds to the *gain term*

$$Q_K^{(+)}[f_1,\dots,f_K](v_1) = K \int_{\mathbb{S}_K \times \mathbb{R}^{dK}} (\otimes_{\ell=1}^K f_\ell)(T_K^\omega V_K)\ \mathrm{d}b_K(\omega)\mathrm{d}v_2 \cdots \mathrm{d}v_K. \tag{7.7}$$

Consequently, f satisfies the equation

$$\partial_t f + \alpha f = \sum_{K=1}^{M} \beta_K Q_K^{(+)}[f,\dots,f], \tag{7.8}$$

where $\alpha = \sum_{K=1}^M \beta_K K > 0$. Thus, Duhamel's formula implies that

$$f(t) = e^{-\alpha t} f_0 + \int_0^t e^{-\alpha(t-s)} \sum_{K=1}^{M} \beta_K Q_K^{(+)}[f(s),\dots,f(s)]\mathrm{d}s, \qquad t \in (0,T_*). \tag{7.9}$$

Next, we adapt the main ideas of the authors in [4][1] and give only a sketch of the proofs. Indeed, we consider the sequence of Picard iterates $\{f_n\}_{n\ge0}$ defined as

$$f_0(t) := f_0, \tag{7.10}$$

$$f_{n+1}(t) := e^{-\alpha t} f_0 + \int_0^t e^{-\alpha(t-s)} \sum_{K=1}^{M} \beta_K Q_K^{(+)}[f_n(s),\dots,f_n(s)]\mathrm{d}s, \qquad n \ge 1. \tag{7.11}$$

[1]We note that the authors consider Picard iterates for an equation similar to ours but in Fourier space and in different functional spaces.

We will use the iterates to show that $\|f(t)\|_{L^1} \le 1$. Indeed, if $\|f_0\|_{L^1} \le 1$, an induction argument shows that $\|f_n(t)\|_{L^1} \le 1$ for all $n \ge 0$ and $t \in (0,T_*)$. Next, note that the gain operators $Q_K^{(+)}$ satisfy the estimate (7.3) with a possibly different constant. Consequently, for a possibly smaller T_*, the following contraction estimate is satisfied thanks to (7.3):

$$\sup_{t\in[0,T_*]} \|f_{n+1}(t) - f_n(t)\|_{L^1} \le \lambda \sup_{t\in[0,T_*]} \|f_n(t) - f_{n-1}(t)\|_{L^1}, \qquad n \ge 1, \tag{7.12}$$

for some fixed $\lambda \in (0,1)$. Thus, the sequence f_n converges to the (unique) solution of (7.9). We conclude that $\|f(t)\|_{L^1} = \lim_{n\to\infty} \|f_n(t)\|_{L^1} \le 1$ for all $t \in (0,T_*)$. This finishes the proof. ∎

Appendix A. Markov processes.

A.1. Review of the general theory. We give a brief review of the basic notions and results from the theory of Markov processes that we use to construct our model; we follow closely the discussion in [22, Chapter 4]. In what follows, we let $(\Sigma, \mathscr{F}, \mathbb{P})$ be a probability space and E a locally compact metric space, with its Borel sets $\mathscr{B}(E)$.
Continuous time. Let us define what we understand for a (continuous-time) Markov process.

Definition 8. *A stochastic process $\boldsymbol{X} = (X(t))_{t=0}^{\infty} : \Sigma \times [0,\infty) \to E$ is called a Markov process if*

$$\mathbb{P}\Big(X(t+s) \in B \,\big|\, \mathscr{F}_t^X\Big) = \mathbb{P}\Big(X(t+s) \in B \,\big|\, \sigma\big(X(t)\big)\Big) \qquad \forall t,s \ge 0,\ \forall B \in \mathscr{B}(E), \tag{A.1}$$

where $\mathscr{F}_t^{\boldsymbol{X}} = \sigma\big(X(s) : 0 \le s \le t\big)$.

Some Markov processes are characterized by more tractable objects. Indeed, let $(T(t))_{t\ge0}$ be a semigroup on $C_b(E)$, the bounded real-valued continuous functions on E.

Definition 9. *We say that the Markov process $\boldsymbol{X}$ corresponds to $(T(t))_{t\ge0}$ if*

$$\mathbb{E}\Big[\varphi\big(X(t+s)\big) \big| \mathscr{F}_t^X\Big] = \big(T(s)\varphi\big)\big(X(t)\big) \qquad \forall t,s \ge 0,\ \forall \varphi \in C_b(E). \tag{A.2}$$

If a Markov process corresponds to a semigroup, it is completely determined by it in the following sense.

Proposition A.1 (see [22, Chapter 4, Proposition 1.6]). *Let $\boldsymbol{X}$ be a Markov process that corresponds to $(T(t))_{t\ge0}$. Then the finite dimensional distributions of $\boldsymbol{X}$ are completely determined by $(T(t))_{t\ge0}$ and the law of $X(0)$.*

Discrete time. Let us define what we understand as a (discrete-time) Markov chain.

Definition 10. *A discrete-time stochastic process $\boldsymbol{Y} = (Y(k))_{k\in\mathbb{N}_0} : \Sigma \times \mathbb{N}_0 \to E$ is called a Markov chain if*

$$\mathbb{P}\Big(Y(n+k) \in B \,\big|\, \mathscr{F}_n^Y\Big) = \mathbb{P}\Big(Y(n+k) \in A \,\big|\, \sigma\big(Y(n)\big)\Big) \qquad \forall n,k \in \mathbb{N}_0,\ \forall B \in \mathscr{B}(E), \tag{A.3}$$

where $\mathscr{F}_n^{\boldsymbol{Y}} = \sigma\big(Y(k) : k \in \{0,\dots,n\}\big)$.

Similarly as before, we can specify Markov chains in terms of more concrete objects. To this end, we define transition functions.

DEFINITION 11. *A function* $\mu : E\times\mathscr{B}(E)\to[0,\infty)$ *is called a transition function if*

$$\begin{cases} \mu(x,\cdot)\in Prob(E) & \forall x\in E,\\ \mu(\cdot,B)\in L^\infty(E) & \forall B\in\mathscr{B}(E). \end{cases} \tag{A.4}$$

DEFINITION 12. *We say that the Markov chain* $\boldsymbol{Y}$ *has* μ *as a transition function if*

$$\mathbb{P}\Big(Y(n+k)\in B\,\big|\,\mathscr{F}_n^Y\Big)=\mu\big(Y(n),B\big)\,,\qquad n\in\mathbb{N}_0,\ B\in\mathscr{B}(E). \tag{A.5}$$

Heuristically, transition functions correspond to the probabilities for the Markov chain to go from one state to the next one. That one may always construct Markov chains with prescribed transition functions and initial laws is the content of the following result.

PROPOSITION A.2 (see [22, Chapter 4, Theorem 1.1]). *For every transition function* μ *and probability measure* $\nu\in Prob(E)$*, there exists a Markov chain* $\boldsymbol{Y}$ *that has* μ *as a transition function and* ν *as the law of* $Y(0)$*.*

Jump processes. If one is given a transition function μ, one may construct Markov processes with explicit transition semigroups; these are called jump processes, which we describe below.

DEFINITION 13. *Let* $\boldsymbol{Y}$ *be a Markov chain with transition function* μ*. We define its generator to be the linear map* $P:C(E)\to C(E)$ *given by*

$$(P\varphi)(x):=\int_E\varphi(y)\mu(x,dy),\qquad x\in E,\ \varphi\in C_b(E)\,. \tag{A.6}$$

Let $(M(t))_{t=0}^\infty$ denote a Poisson process with parameter λ independent of $\mathbf{Y}$. We let the *jump process* associated to $\mathbf{Y}$ with parameter λ be the stochastic process $\mathbf{V}=(V(t))_{t=0}^\infty$ defined by

$$V(t):=Y\big(M(t)\big),\qquad t\ge 0\,. \tag{A.7}$$

PROPOSITION A.3 (see [22, Chapter 4, section 2]). *The stochastic process* $\boldsymbol{V}$ *defined by* (A.7) *is a Markov process that corresponds to the semigroup* $\{\exp\big(t\,\lambda(P-Id)\big)\}_{t\ge0}$.

Here we will give a sketch of the proof of the above proposition.

Proof. Let $\phi\in C(E)$. The transition semigroup $T(t)$ for the Markov process $V(t)$ is defined through $T(s)\phi\big(V(t)\big)=\mathbb{E}[\phi\big(V(t+s)\big)|\mathscr{F}_t]$. Using the memoryless property of the Poisson process $M(t)$ along with the law of total probability, we can calculate

$$\begin{aligned} T(s)\phi\big(V(t)\big)&=\mathbb{E}[\phi\big(V(t+s)\big)|\mathscr{F}_t]=\mathbb{E}[\phi\Big(Y\big(M(t+s)\big)\Big)|\mathscr{F}_t]\\ &=\mathbb{E}[\phi\Big(Y\big(M(t+s)-M(t)+M(t)\big)\Big)|\mathscr{F}_t]\\ &=\sum_{k\ge0}\mathbb{P}\big(M(t+s)-M(t)=k\big)\mathbb{E}[\phi\Big(Y\big(k+M_t\big)\Big)|\mathscr{F}_t]\\ &=\sum_{k\ge0}e^{-\lambda s}\frac{(\lambda s)^k}{k!}P^k\phi\big(V(t)\big)\\ &=\exp\big(t\lambda(P-\mathrm{Id})\big)\phi\big(V(t)\big). \end{aligned} \tag{A.8}$$

This finishes the sketch of the proof. ∎

DEFINITION 14. *A Poisson process $M(t)$ with rate $\lambda > 0$ is a stochastic process taking values on $\mathbb{N}$ with the following conditions:*

1. $M(0) = 0$.
2. *For all $s_i < t_i$, the increments $M(t_i) - M(s_i)$ are independent random variables.*
3. $\mathbb{E}[M(t)] = \lambda t$.

Furthermore, the Poisson process is a Markov process and thus has the "memoryless" property, implying that its increments satisfy,

$$M(t_i) - M(s_i) = M(t_i - s_i) \quad \forall s_i < t_i. \tag{A.9}$$

Acknowledgments. N.P. would like to thank Sergio Simonella and Mario Pulvirenti for discussions that inspired this research direction. The authors would like to thank Joseph K. Miller and Chiara Saffirio for helpful discussions. Also, the authors would like to thank the anonymous referees for their constructive comments that helped improve the paper.

REFERENCES

[1] I. Ampatzoglou and N. Pavlović, *Rigorous derivation of a ternary Boltzmann equation for a classical system of particles*, Comm. Math. Phys., 387 (2021), pp. 793–863.

[2] I. Ampatzoglou and N. Pavlović, *Rigorous Derivation of a Binary-Ternary Boltzmann Equation for a Dense Gas of Hard Spheres*, preprint, https://arxiv.org/abs/2007.00446, 2020.

[3] I. Ampatzoglou, I. M. Gamba, N. Pavlović, and M. Tasković, *Moment Estimates and Well-Posedness of the Binary-Ternary Boltzmann Equation*, preprint, https://arxiv.org/abs/2210.09600, 2022.

[4] A. V. Bobylev, C. Cercignani, and I. M. Gamba, *On the self-similar asymptotics for generalized non-linear kinetic Maxwell molecules*, Comm. Math. Phys., 291 (2009), pp. 599–644.

[5] L. Boltzmann, *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*, Sitz.-Ber. Akad. Wiss. Wien (II), 66 (1872), pp. 275–370.

[6] F. Bonetto, A. Geisinger, M. Loss, and T. Ried, *Entropy decay for the Kac evolution*, Comm. Math. Phys., 363 (2018), pp. 847–875.

[7] F. Bonetto, M. Loss, and R. Vaidyanathan, *The Kac model coupled to a thermostat*, J. Stat. Phys., 156 (2013), pp. 647–667.

[8] E. A. Carlen, M. C. Carvalho, and M. Loss, *Chaos, ergodicity and equilibria in a quantum Kac model*, Adv. Math., 358 (2019), 106827.

[9] E. A. Carlen, M. C. Carvalho, J. Le Roux, M. Loss, and C. Villani, *Entropy and chaos in the Kac model*, Kinet. Relat. Models, 3 (2010), pp. 85–122.

[10] E. A. Carlen, M. C. Carvalho, and M. Loss, *Many-body aspects of approach to equilibrium*, in Seminaire: Équations aux Dérivées Partielles, 2000–2001, Śemine Equ. Dériv. Partielles, pages Exp. No. XIX, 12, Ecole Polytech., Palaiseau, 2001.

[11] E. A. Carlen, M. C. Carvalho, and M. Loss, *Determination of the spectral gap for Kac's master equation and related stochastic evolution*, Acta Math., 191 (2003), pp. 1–54.

[12] E. A. Carlen, J. S. Geronimo, and M. Loss, *Determination of the spectral gap in the Kac Model for physical momentum and energy-conserving collisions*, SIAM J. Math. Anal., 40 (2008), pp. 327–364, https://doi.org/10.1137/070695423.

[13] E. Carlen, D. Mustafa, and B. Wennberg, *Propagation of chaos for the thermostated Kac master equation*, J. Stat. Phys., 158 (2015), pp. 1341–1378.

[14] K. Carrapatoso, *Quantitative and qualitative Kac's chaos on the Boltzmann's sphere*, Ann. Inst. Henri Poincaré Probab. Stat., 51 (2015), pp. 993–1039.

[15] C. Cercignani, *The Theory and Application of the Boltzmann Equation*, Appl. Math. Sci. 67, Springer, New York, 1988.

[16] R. Cortez, *Uniform propagation of chaos for Kac's 1D particle system*, J. Stat. Phys., 165 (2016), pp. 1102–1113.

[17] R. Cortez and H. Tossounian, *Uniform propagation of chaos for the thermostated Kac model*, J. Stat. Phys., 183 (2021), 28.

[18] R. Cortez and J. Fontbona, *Quantitative uniform propagation of Chaos for Maxwell molecules*, Comm. Math. Phys., 357 (2018), pp. 913–941.

[19] P. Diaconis and L. Saloff-Coste, *Bounds for Kac's master equation*, Comm. Math. Phys., 209 (2000), pp. 729–755.

[20] A. Einav, *On Villani's conjecture concerning entropy production for the Kac master equation*, Kinet. Relat. Models, 4 (2011), pp. 479–497.

[21] L. Erdös, B. Schlein, and H.-T. Yau, *Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems*, Invent. Math., 167 (2007), pp. 515–614.

[22] S. Ethier and T. Kurtz, *Markov Processes*, Wiley Interscience, New York, 1986.

[23] I. Gallagher, L. Saint-Raymond, and B. Texier, *From Newton to Boltzmann: Hard Spheres and Short-Range Potentials*, Zur. Lect. Adv. Math., European Mathematical Society, Zurich, 2013.

[24] C. Graham and S. Méléard, *Stochastic particle approximations for generalized Boltzmann models and convergence estimates*, Ann. Probab., 25 (1997), pp. 115–132.

[25] M. Hauray, *Uniform contractivity in Wasserstein metric for the original* 1*D Kac's model*, J. Stat. Phys., 162 (2016), pp. 1566–1570.

[26] M. Hauray and S. Mischler, *On Kac's chaos and related problems*, J. Funct. Anal., 266 (2014), pp. 6055–6157.

[27] E. Janvresse, *Spectral gap for Kac's model of Boltzmann equation*, Ann. Probab., 29 (2001), pp. 288–304.

[28] M. Kac, *Foundations of kinetic theory*, in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Contributions to Astronomy and Physics 3, University of California Press, Berkeley, CA, 1956, pp. 171–197.

[29] O. E. Lanford, *Time evolution of large classical systems*, in Dynamical Systems, Theory and Applications, Lecture Notes in Phys. 38, J. Moser, ed., Springer, Berlin, Heidelberg, 1975.

[30] J. Maxwell, *On the dynamical theory of gases*, Philos. Trans. Roy. Soc. London Ser. A, 157 (1867), pp. 49–88.

[31] D. Maslen, *The eigenvalues of Kac's master equation*, Math. Z., 243 (2003), pp. 291–331.

[32] H. P. Mckean, Jr., *An exponential formula for solving Boltzmann's equation for a Maxwellian gas*, J. Combin. Theory, 2 (1967), pp. 358–382.

[33] S. Mischler and C. Mouhot, *Kac's program in kinetic theory*, Invent. Math., 193 (2013), pp. 1–147.

[34] S. Mischler, C. Mouhot, and B. Wennberg, *A new approach to quantitative chaos propagation for drift, diffusion and jump processes*, Probab. Theory Related Fields, 161 (2015), pp. 1–59.

[35] H. Tanaka, *Propagation of chaos for certain Markov processes of jump type with nonlinear generators*, I, Proc. Japan Acad., 45 (1969), pp. 449–452.

[36] H. Tanaka, *Propagation of chaos for certain Markov processes of Jump type with nonlinear generators*, II, Proc. Japan Acad., 45 (1969), pp. 598–600.

[37] H. Tanaka, *Probabilistic treatment of the Boltzmann equation of Maxwellian molecules*, Z. Wahrsch. Verw. Gebiete, 46 (1978), pp. 67–105.

[38] T. Ueno, *A class of Markov processes with non-linear, bounded generators*, Jpn. J. Math., 38 (1969), pp. 19–38.

[39] C. Villani, *Cercignani's conjecture is sometimes true and always almost true*, Comm. Math. Phys., 234 (2003), pp. 455–490.