Markov Processes Relat. Fields **12**, 95–138 (2006)

Markov Processes and Related Fields

Kinetic Equations for the Pure Jump Models of *k*-nary Interacting Particle Systems

V.N. Kolokoltsov

Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoy Karetny per., 19, Moscow, Russia and Centro de Investigacion en Matematicas (CIMAT), Guanajuato, CP 36000, Mexico. E-mail: vkolok@fsmail.net

Received December 8, 2004, revised September 26, 2005

Abstract. The dynamic law of large numbers of k-nary interacting particles is described by nonlinear measure-valued equations that generalize the classical kinetic equations of statistical mechanics and the evolutionary (replicator) dynamics from population biology and evolutionary games. Rigorous results on convergence, existence and uniqueness theorems for the limiting equations, and the propagation of chaos property are established for models with interactions of pure jump type.

KEYWORDS: interacting particles, k-nary interaction, measure-valued limits, kinetic equation, mean-field limit, mass exchange processes, coagulation-fragmentation, propagation of chaos, evolutionary games, replicator dynamics

AMS SUBJECT CLASSIFICATION: 60K35, 82C21, 82C22

1. Introduction

1.1. Aims of the paper

This paper presents a continuation of the author's work on the dynamic law of large numbers for k-nary interacting particle systems, see [19, 20], and extends the results of [21] devoted to mass exchange processes with discrete mass distribution to the models with a general state space. The aims of the paper are the following: (i) by combining the ideas of the mean-field (or McKean–Vlasov) and k-nary interaction, to show that the general deterministic law of large numbers for Markovian models of interacting particles is described by a non-linear positivity preserving measure-valued equation, equation (3.12), that includes as particular cases all kinetic equations of statistical mechanics (Smoluchovski, Landau, Boltzmann, Vlasov), their combinations and extensions for k-nary interacting particles derived in [4] (including non-binary coagulations or fragmentation, where, say, the rate of coagulation or fragmentation of two particles can be increased or decreased by the presence of a third particle) and general evolutionary dynamics (or replicator dynamics) from population biology and evolutionary games; (ii) for the case of particle systems with a pure jump interaction, to prove existence and uniqueness results for the Cauchy problem of limiting equations, equations (3.10), and the weak convergence of the approximating Markov models of interacting particles to the solutions of these equations.

We discuss here only deterministic laws of large numbers. A natural extension of our limiting procedure (beyond the scope of the present paper) leads to measure-valued stochastic limits of interacting particle systems. For the case of finite number of particles these general stochastic limits were analyzed in [19, 20], see also [22, 24] for a necessary background on boundary value problems for pseudo-differential operators with growing symbols and [21] for an example of the diffusion limit of a mass exchange process. Roughly speaking, the general class of limiting stochastic processes represents a subclass of measure-valued Feller processes having (infinite-dimensional) pseudo-differential generators with symbols depending polynomially on the position (for instance, linear dependence corresponds to super-processes).

1.2. Content of the paper

In the next section, we shall fix some notations used throughout the paper, in particular those connected with the state space of a system of interacting particles, and recall the basic notions and facts from the theory of pure jump Markov processes. In Section 3 we describe the general Markov processes that model the k-nary interacting exchangeable particle systems and give a heuristic derivation of the corresponding kinetic equations, equations (3.12), discussing also various forms of these equations and their main examples. In particular, we show that by including special mean-field interaction (that leads effectively to a different scaling of k-nary interactions than the one mostly used in statistical mechanics models) our model yields the far reaching generalizations (with k-nary interaction specified by arbitrary generators and state spaces) of the replicator dynamics models from population biology and evolutionary games. Inspired by biological analogs we propose in passing a new type of coagulation-fragmentation equations obtained by scaling the k-nary interactions in a way used in population biology.

In Section 4, we justify the heuristic arguments of the previous section for pure jump Markov models proving (subject to some additional assumptions) that a subsequence of the family of pure jump Markov processes describing our k-nary interacting exchangeable particle systems converges to the deterministic measure-valued process defined by the integral version of the weak kinetic equations. In the literature on coagulation-fragmentation one can distinguish two approaches to proving existence of solutions to kinetic equations: probabilistic (by a weak limit from approximating particle systems, see e.g. [30, 31]; this approach being difficult to follow for non-probabilists, as it requires the heavy modern tools of the theory of stochastic processes) and deterministic (by approximating the transition kernel by bounded ones and then passing to the limit in the equation, see e.g. [2,13,26] and references therein). We develop both approaches unifying in passing a variety of known results. Ideologically the probabilistic approach seems to be more relevant, as it supplies in passing a rigorous connection between limiting kinetic equations and the original approximating particle system. But deterministic approach has the advantage of usually being simpler and at the same time requiring weaker assumptions on the model. This is reflected in our exposition, where we first obtain several existence results probabilistically, and then obtain the improvement of some of them deterministically. An application of our results to binary interactions yields already seemingly new rigorous result on the convergence of natural stochastic approximations to the solutions of the kinetic equations of the standard coagulation-fragmentation model filling the gap between the corresponding results from [30, 31] (pure coagulation model) and [18] (coagulation-fragmentation with discrete mass distribution). For replicator dynamics of evolutionary biology our result seem to be completely new in the general measure-valued settings. We conclude this section by an important regularity result showing that our weak solutions turn out to be also strong solutions under mild additional assumptions. This result seems to be new even for a standard coagulation-fragmentation model.

Section 5 is devoted to a slightly different approximation of kinetic equations, namely to an approximation by Markov chains obtained by the discretization of the state space X. This approximation seems to be of interest for applications, say, for computer simulations of our models, and it can be considered as a stochastic counterpart of deterministic discretizations of kinetic models (for the latter see [28] and references therein).

In Section 6 we obtain the well-posedness of the Cauchy problem for our kinetic equations by proving uniqueness and continuous dependence of the solutions on the initial conditions under some reasonable assumptions. Even for binary coagulation-fragmentation models, this theorem improves the known results on uniqueness obtained in [13], where the uniqueness is proved only under a restrictive assumption of the existence of finite exponential moments (and for measures with a density with respect to Lebesgue measure), and on the other hand, the results on uniqueness from [30, 31] obtained there for a binary pure coagulation model. Our method of proving uniqueness can be considered as a non-trivial extension of the method of proving uniqueness developed for discrete coagulation type models in [2, 21, 29] (that effectively goes back to Povsner's method [32] on the analysis of Boltzmann equation). This extension is based on a measure theoretic result presented as Lemma A in the Appendix. As a corollary of our main result, we prove for conclusion a version of the propagation of chaos property for our model of k-nary interacting particle systems (cf. [35] for the corresponding result for the Boltzmann equation). As another corollary, let us stress that our uniqueness result essentially enlarges the class of initial distributions for which the discrete approximation scheme discussed in [28] converges to the solution of the fragmentation-coagulation equation.

The paper ends with Appendix, where a measure-theoretical Lemma A is proved (that is needed for the proof of uniqueness) and a simple differential equation is solved in Lemma B (that is needed to get crucial lower bounds for the solutions of our kinetic equations).

2. Preliminaries

2.1. Some notations

We list here a few notations that will be used throughout the paper without further reminder:

> $\mathbf{1}$ denotes the function that equals identically 1 or the identity operator in a Banach space; $\mathbf{1}_M$ for a set M denotes the indicator function of M that equals 1 for $x \in M$ and vanishes otherwise; $o(1)_{x\to a}$ denotes a function depending on x that tends to zero as $x \to a$;
>
> for a measurable space Y, $B(Y)$ denotes the Banach space of real bounded measurable functions on Y equipped with the usual sup-norm; if Y is a topological space, $C_b(Y)$ denotes the Banach subspace of $B(Y)$ consisting of continuous functions; $\mathcal{M}(Y)$ is the space of finite (signed) measures on Y, considered as a Banach space with the norm $\|\cdot\|$ on $\mathcal{M}(Y)$ being the total variation norm; for $\mu \in \mathcal{M}(Y)$ we shall denote by $|\mu|$ the total variation measure of μ so that $\|\mu\| = \int_Y |\mu|(dy)$;
>
> if X is a locally compact space, then $C_0(X)$ (respectively $C_c(X)$) is the Banach space of continuous functions vanishing at infinity and equipped with sup-norm (respectively its subspace consisting of functions with a compact support);
>
> the upper subscript "+" for all these spaces (e.g. $C_0^+(X)$, $\mathcal{M}^+(X)$) will denote the corresponding cones of non-negative elements;
>
> by a symmetric function of n (vector) variables we shall understand a function, which is symmetric with respect to all permutations of these variables, and by a symmetric operator on the space of functions of n

variables we shall understand an operator that preserves the set of symmetric functions;

for a finite subset $I = \{i_1, \dots, i_k\}$ of a countable set J, we denote by $|I|$ the number of elements in I, by $\bar{I}$ its complement $J \setminus I$, by $\mathbf{x}_I$ the collection of the variables $x_{i_1}, \dots, x_{i_k}$ and by $d\mathbf{x}_I$ the measure $dx_{i_1} \dots dx_{i_k}$.

A transitional kernel from X to Y means, as usual, a measurable function $\mu(x, \cdot)$ from $x \in X$ to the cone of positive finite measures on Y.

We shall denote as usual by $D_X[0, \infty)$ the Skorokhod space of càdlàg paths $[0, \infty) \mapsto X$ equipped with the standard filtration $\mathcal{F}_t$. We shall denote by P and E the probability and the expectation of events and functions.

2.2. State space and observables of systems of interacting particles

Throughout the paper we shall denote by X a locally compact metric space equipped with its Borel sigma algebra. Denoting by X^0 a one-point space and by X^j the powers $X \times \dots \times X$ (j-times) considered with their product topologies, we shall denote by $\mathcal{X}$ their disjoint union $\mathcal{X} = \cup_{j=0}^{\infty} X^j$, which is again a metric locally compact space. In applications, X specifies the state space of one particle and $\mathcal{X} = \cup_{j=0}^{\infty} X^j$ stands for the state space of a random number of similar particles. We shall denote by $B_{sym}(\mathcal{X})$ (resp. $C_{sym}(\mathcal{X})$) the Banach spaces of symmetric bounded measurable (resp. continuous) functions on $\mathcal{X}$ and by $B_{sym}(X^k)$ (resp. $C_{sym}(X^k)$) the corresponding spaces of functions on the finite power X^k. The space of symmetric measures will be denoted by $\mathcal{M}_{sym}(\mathcal{X})$. The elements of $\mathcal{M}_{sym}^+(\mathcal{X})$ and $C_{sym}(\mathcal{X})$ are called respectively the (mixed) states and observables for a Markov process on $\mathcal{X}$. We shall denote the elements of $\mathcal{X}$ by bold letters, e.g. $\mathbf{x}$, $\mathbf{y}$. Sometimes it is convenient to consider the factor spaces SX^k and $S\mathcal{X}$ obtained by the factorization of X^k and $\mathcal{X}$ with respect to all permutations, which allows for the identifications $C_{sym}(\mathcal{X}) = C(S\mathcal{X})$ and likewise. Symmetrical laws on X^k (which are uniquely defined by their projections to SX^k) are called exchangeable systems of k particles. A key observation for the theory of measure-valued limits is the inclusion $S\mathcal{X}$ to $\mathcal{M}^+(X)$ given by

$$\mathbf{x} = (x_1, \dots, x_l) \mapsto \delta_{x_1} + \dots + \delta_{x_l}, \tag{2.1}$$

which defines a bijection between $S\mathcal{X}$ and the space $\mathcal{M}_\delta^+(X)$ of finite linear combinations of δ-measures.

Clearly each $f \in B_{sym}(\mathcal{X})$ is defined by its components f^k on X^k so that for $\mathbf{x} = (x_1, \dots, x_k) \in X^k \subset \mathcal{X}$, say, one can write $f(\mathbf{x}) = f(x_1, \dots, x_k) = f^k(x_1, \dots, x_k)$ (the upper index k at f is optional and is used to stress the number of variables in an expression). Similar notations are for measures. In particular, the pairing between $C_{sym}(\mathcal{X})$ and $\mathcal{M}(\mathcal{X})$ can be written as

$$(f,\rho) = \int f(\mathbf{x})\rho(d\mathbf{x}) = f^0\rho_0 + \sum_{n=1}^{\infty} \int f(x_1,\ldots,x_n)\,\rho(dx_1\ldots dx_n),$$
$$f \in C_{sym}(\mathcal{X}),\ \ \rho \in \mathcal{M}(\mathcal{X}), \qquad (2.2)$$

so that $\|\rho\| = (\mathbf{1},\rho)$ for $\rho \in \mathcal{M}^+(\mathcal{X})$.

A useful class of measures (and mixed states) on $\mathcal{X}$ is given by the decomposable measures of the form $Y^{\otimes}$ and $Y^{\tilde{\otimes}}$, which are defined for an arbitrary finite measure $Y(dx)$ on X by their components

$$(Y^{\otimes})_n(dx_1\ldots dx_n) = Y^{\otimes n}(dx_1\ldots dx_n) = Y(dx_1)\ldots Y(dx_n) \qquad (2.3)$$

and

$$(Y^{\tilde{\otimes}})_n(dx_1\ldots dx_n) = \frac{1}{n!}Y^{\otimes n}(dx_1\ldots dx_n) = \frac{1}{n!}Y(dx_1)\ldots Y(dx_n). \qquad (2.4)$$

Notice that unlike $Y^{\tilde{\otimes}}$ the measure $Y^{\otimes}$ needs not to be a finite measure on $\mathcal{X}$ even when Y is finite. Similarly the decomposable observables (or exponential vectors) are defined for an arbitrary $Q \in C(X)$ as

$$(Q^{\otimes})^n(x_1,\ldots,x_n) = Q^{\otimes n}(x_1,\ldots,x_n) = Q(x_1)\ldots Q(x_n). \qquad (2.5)$$

We shall use also the additively decomposable observables defined for an arbitrary $Q \in C(X)$ as

$$(Q^+)(x_1,\ldots,x_n) = Q(x_1) + \ldots + Q(x_n) \qquad (2.6)$$

(Q^+ vanishes on X^0). In particular, if $Q = \mathbf{1}$, then $Q^+ = \mathbf{1}^+$ is the number of particles: $\mathbf{1}^+(x_1,\ldots,x_n) = n$.

We shall deduce now a simple combinatorial identity for symmetric functions that will be used in our derivation of kinetic equations. We shall denote by δ_x the Dirac measure at x. By a Young scheme Γ we mean a collection $\Gamma = \{\gamma_1,\ldots,\gamma_l\}$ of natural numbers such that $1 \le \gamma_1 \le \ldots \le \gamma_l$. Let $h > 0$ be a positive parameter.

Proposition 2.1. *For any natural k, there exist constants α_Γ parametrized by all Young schemes $\Gamma = \{\gamma_1,\ldots,\gamma_l\}$ with $\gamma_1 + \ldots + \gamma_l = k$ such that for any natural n, $f \in B_{sym}(X^k)$ and a collection of points $x_1,\ldots,x_n$ in X*

$$h^k \sum_{I\subset\{1,\ldots,n\},|I|=k} f(x_I) = \frac{1}{k!}\int f(y_1,\ldots,y_k)\prod_{j=1}^{k}(h\delta_{x_1} + \ldots + h\delta_{x_n})(y_j)$$
$$+ \sum_{\Gamma} \alpha_\Gamma h^{k-l} \int f(y_1,\ldots,y_1,y_2,\ldots,y_2,\ldots,y_l,\ldots,y_l)$$

$$\times \prod_{j=1}^{l} (h\delta_{x_1} + \ldots + h\delta_{x_n})(y_j), \tag{2.7}$$

where $\sum_\Gamma$ *is the sum over all Young schemes* $\Gamma = \{\gamma_1, \ldots, \gamma_l\}$ *such that* $\gamma_1 + \ldots + \gamma_l = k$ *and* $\gamma_l > 1$ *(or, equivalently,* $l < k$*), and notation* $f(y_1, \ldots, y_1, y_2, \ldots \ldots, y_2, \ldots, y_l, \ldots, y_l)$ *means that the first* γ_1 *arguments of* f *are equal to* y_1*, the next* γ_2 *arguments are equal to* y_2*, etc. Moreover, if* f *is non-negative, then the l.h.s. of* (2.7) *does not exceed the first term of the r.h.s. of* (2.7).

Proof. The proof is obtained by trivial induction in k from the following evident formula

$$h^k \sum_{i_1, \ldots i_k = 1}^{n} f(x_{i_1}, \ldots, x_{i_k}) = \int f(y_1, \ldots y_k) \prod_{j=1}^{k} (h\delta_{x_1} + \ldots + h\delta_{x_n})(y_j). \tag{2.8}$$

To avoid long expressions, we shall just deduce (2.7) for $k = 2$ and $k = 3$. We have

$$\begin{aligned}
h^2 \sum_{I \subset \{1, \ldots, n\}, |I| = 2} f(x_I) &= \frac{h^2}{2} \sum_{i_1 \neq i_2} f(x_{i_1}, x_{i_2}) \\
&= \frac{h^2}{2} \Big[\sum_{i_1, i_2 = 1}^{n} f(x_{i_1}, x_{i_2}) - \sum_{i=1}^{n} f(x_i, x_i) \Big] \\
&= \frac{1}{2} \int f(y_1, y_2) \prod_{j=1}^{2} (h\delta_{x_1} + \ldots + h\delta_{x_n})(y_j) \\
&\quad - \frac{h}{2} \int f(y, y)(h\delta_{x_1} + \ldots + h\delta_{x_n})(y).
\end{aligned}$$

Similarly,

$$\begin{aligned}
h^3 \sum_{I \subset \{1, \ldots, n\}, |I| = 3} f(x_I) &= \frac{h^3}{3!} \sum_{i_1, i_2, i_3 : i_1 \neq i_2, i_1 \neq i_3, i_2 \neq i_3} f(x_{i_1}, x_{i_2}, x_{i_3}) \\
&= \frac{h^3}{3!} \sum_{i_1, i_2, i_3} f(x_{i_1}, x_{i_2}, x_{i_3}) \\
&\quad - \frac{h^3}{2} \sum_{i_1 \neq i_2} f(x_{i_1}, x_{i_1}, x_{i_2}) - \frac{h^3}{3!} \sum_i f(x_i, x_i, x_i) \\
&= \frac{1}{3!} \int f(y_1, y_2, y_3) \prod_{j=1}^{3} (h\delta_{x_1} + \ldots + h\delta_{x_n})(y_j) \\
&\quad - \frac{h}{2} \int f(y_1, y_1, y_2) \prod_{j=1}^{2} (h\delta_{x_1} + \ldots + h\delta_{x_n})(y_j)
\end{aligned}$$

$$+\frac{h^2}{3}\int f(y,y,y)(h\delta_{x_1}+\ldots+h\delta_{x_n})(y).$$

The last assertion of the proposition is evident, because one sees from the above that, whenever f is non-negative, the first term of the asymptotics (2.7) is obtained by increasing the sum from the l.h.s. of (2.7). □

2.3. Basic facts about jump processes

By a pure jump Markov process on X we mean a Markov process with a generator of the form

$$(Gf)(x)=\int(f(y)-f(x))q(x,dy)=\int f(y)q(x;dy)-f(x)q(x), \tag{2.9}$$

where $(q(x),q(x,\cdot))$ is a totally stable conservative q-pair, which means that (see e.g. [8]) $q(x,\cdot)$ is a transition kernel from X to X such that $q(x,\{x\})=0$, and $q(x)=q(x,X)$ called the intensity of the q-pair, is everywhere finite. An important question is whether one can construct a Markov process with a given generator of type (2.9) and whether such a process is unique. We refer to [8] and [15] for the general theory. We shall need the following corollary of this theory, which proof we sketch for completeness.

Proposition 2.2. (i) *Suppose $x\mapsto q(x,\cdot)$ is a uniformly bounded function $X\mapsto\mathcal{M}^+(X)$ that is continuous with respect to the weak topology of $\mathcal{M}^+(X)$. Then for any $x\in X$ there exists a unique Markov process Z_t^x on X with generator* (2.9)*, whose domain is the whole Banach space $C_b(X)$, and moreover Z_t^x can be characterized as the unique solution to the corresponding martingale problem, i.e. as the unique distribution on the Skorohod space $D_X[0,\infty)$ such that*

$$M_t=f(Z_t^x)-f(x)-\int_0^t Gf(Z_s^x)\,ds \tag{2.10}$$

is an $\mathcal{F}_t$-martingale for all $f\in C_c(X)$ such that $M_0=0$ almost surely. The corresponding semigroup $T_tf(x)=\mathsf{E}\,f(Z_t^x)$ is a strongly continuous semigroup of contractions on $C_b(X)$ and (2.10) *is a martingale for any non-negative continuous f on X such that $Gf(x)\le b+cf(x)$ for all x and some constants $b,c\ge 0$.*

(ii) *Suppose again that $x\mapsto q(x,\cdot)$ is weakly continuous, but may be not bounded, and that there exists a continuous non-negative function ψ on X such that the intensity $q(x)$ is uniformly bounded on all sets $U_a=\{x:\psi(x)\le a\}$, and moreover $G\psi(x)\le b+c\psi(x)$ for all x and some constants $b,c\ge 0$. Then there exists a unique Markov process Z_t^x on X starting at x having a generator that is an extension of* (2.9) *considered on the domain $C_c(X)$, and moreover Z_t^x*

can be characterized as the unique distribution on the Skorohod space $D_X[0,\infty)$ such that (2.10) *is a martingale for all $f \in C_c(X)$. The corresponding semigroup $T_t f(x) = \mathsf{E} f(Z_t^x)$ is a semigroup of contractions on $C_b(X)$ (but not necessarily strongly continuous). At last, for any positive T and r*

$$\mathsf{P}\big(\sup_{t\in[0,T]} \psi(Z_t^x) > r\big) \le \frac{C(T)}{r}, \tag{2.11}$$

where $C(T)$ depends only on T,b and $\psi(x)$ (and not on the bounds of the intensity q on U_a.

Sketch of the proof. (i) In this case, G is a bounded operator on $C_b(X)$ and the semigroup is uniquely obtained as $T_t = e^{Gt}$ (defined as a converging series). This implies the well-posedness of the martingale problem for G with the domain C_b (see e.g. Theorem 4.1 of Chapter 4 from [15]). To get the well-posedness of the martingale problem for G with the domain C_c one can, for example, approximate G by operators G_n obtained by multiplication of $q(x,\cdot)$ by continuous functions $X \mapsto [0,1]$ with compact support (such G_n define a strongly continuous semigroup in $C_0(X)$ with a core C_c and consequently the martingale problem is well posed for G_n defined on $C_c(X)$), and then apply the localization theory (see e.g. Theorem 6.2 of Chapter 4 from [15]). At last, let a continuous f be non-negative and $Gf \le b+cf$. Let $f_n(x) = \min(n, f(x))$. Then $Gf_n \le b + cf_n$ for all n (in particular, $Gf_n(x) \le 0$ whenever $f_n(x) = n$). As $0 \le f_n \le n$, (2.10) is a martingale for f_n playing the role of f. By Gronwall's lemma this implies

$$\mathsf{E} f_n(Z_t^x) \le (f(x) + tb)e^{ct}.$$

By the monotone convergence theorem this implies the same estimate for f on the place of f_n. Hence approximating f by f_n in (2.10) we can conclude that (2.10) holds by the dominated convergence theorem.

(ii) Instead of G, consider its approximation G_n defined as

$$(G_n f)(x) = \int (f(y) - f(x))\tilde{\mathbf{1}}_{\psi(\cdot)\le n} q(x,dy),$$

where $\tilde{\mathbf{1}}_{\psi(\cdot)\le n}$ is a continuous function $X \mapsto [0,1]$ that coincides with $\mathbf{1}_{\psi(\cdot)\le n}$ on the set of all x where $\psi(x) \notin [n, n+1]$. Each G_n satisfies conditions (i). Moreover, as G_n are obtained from G by the multiplication on $\mathbf{1}_{\psi(x)\le n}$, they enjoy the same property $G_n\psi(x) \le b + c\psi(x)$. Hence

$$\psi(Z_t^{x,n}) - \psi(x) - \int_0^t G_n \psi(Z_s^{x,n})\, ds \tag{2.12}$$

is a martingale and

$$\mathsf{E}\psi(Z_t^{x,n}) \le (\psi(x) + tb)e^{ct}. \tag{2.13}$$

Hence, for any T and with probability arbitrary close to one, the processes $\psi(Z_t^{x,n})$ are uniformly bounded for $t \in [0,T]$, and hence G_n form a localizing sequence for G such that Theorem 6.3 from Chapter 4 of [15] can be applied to get the well-posedness of the martingale problem for G and hence the existence and uniqueness of the corresponding Markov process. At last, passing to the limit as $n \to \infty$ in (2.12), (2.13) yields the same properties for $\psi(Z_t^x)$, and hence using the Doob maximal inequality for martingales yields (2.11).

In the future, we shall be interested in pure jump processes on $\mathcal{X}$, whose semigroup and the generator preserves the space C_{sym} of continuous symmetric functions so that the q-pair is given by symmetric transition kernels $q(\mathbf{x}; d\mathbf{y})$ that could be thus considered as kernels on the factor space $S\mathcal{X}$.

For conclusion, let us notice that though we have chosen to work with locally compact metric spaces X for simplicity and transparency (for the same purpose we also assume the continuity of all kernels), most of the results can be generalized (with minor modifications) to Polish spaces X or even to general measurable spaces X, as considered in [31] for coagulation models.

3. Heuristic derivation of kinetic equations and examples

3.1. Pure jump Markov models for k-nary interacting particles

A k-nary interaction of a pure jump type is specified by a transition kernel

$$P^k(x_1, \dots, x_k; d\mathbf{y}) = \{P^k_m(x_1, \dots, x_k; dy_1 \dots dy_m)\}$$

from SX^k to $S\mathcal{X}$ such that $P^k(\mathbf{x}; \{\mathbf{x}\}) = 0$ for all $\mathbf{x} \in \mathcal{X}$, and with the intensity

$$P^k(x_1, \dots, x_k) = \int P^k(x_1, \dots, x_k; d\mathbf{y}) = \sum_{m=0}^{\infty} \int P^k_m(x_1, \dots, x_k; dy_1 \dots dy_m).$$

The intensity defines the rate of decay of any collection of k particles $x_1, \dots, x_k$ and the measure $P^k(x_1, \dots, x_k; d\mathbf{y})$ defines the distribution of possible outcomes. Supposing that any k particles from a given set of $n \geq k$ particles can interact, we arrive to the following generator of k-nary interacting particles defined by the kernel P^k:

$$\begin{aligned}(G_k f)(x_1, \dots, x_n) &= \sum_{m=0}^{\infty} \sum_{I \subset \{1,\dots,n\}, |I|=k} \int (f^{l+(n-k)}(\mathbf{x}_{\bar{I}}, y_1, \dots, y_m) \\ &\quad - f(x_1, \dots, x_n)) P^k_m(x_I; dy_1 \dots dy_m) \\ &= \sum_{I \subset \{1,\dots,n\}, |I|=k} \int (f_{\mathbf{x}_{\bar{I}}}(\mathbf{y}) - f(x_1, \dots, x_n)) P^k(x_I, d\mathbf{y}),\end{aligned} \tag{3.1}$$

where we have introduced the notation $f_{\mathbf{x}}(\mathbf{y}) = f(\mathbf{x}, \mathbf{y})$. Taking into account all possible interactions of order $\leq k$ given by the kernels $P = \{P^l_m\}$, $l = 0, \ldots, k$, $m = 0, 1, \ldots$, from $S\mathcal{X}$ to $S\mathcal{X}$, with the intensity

$$P(\mathbf{x}) = \int P(\mathbf{x}; d\mathbf{y}) = \int \sum_{m=0}^{\infty} P^l_m(x_1, \ldots, x_l; dy_1, \ldots, dy_m) \tag{3.2}$$

whenever $\mathbf{x} = (x_1, \ldots, x_l)$ with $l \leq k$, and $P(\mathbf{x}) = 0$ whenever $\mathbf{x} \in X^l$ with $l > k$, yields the generator of a general pure jump Markov processes with the interaction of order $\leq k$,

$$(G_{\leq k} f)(x_1, \ldots, x_n) = \sum_{I \subset \{1, \ldots, n\}} \int (f_{\mathbf{x}_{\bar{I}}}(\mathbf{y}) - f(x_1, \ldots, x_n)) P(\mathbf{x}_I, d\mathbf{y}). \tag{3.3}$$

Changing the state space by (2.1) yields the corresponding Markov process on $\mathcal{M}^+_\delta(X)$. Choosing a positive parameter h, we shall perform now the following scaling: firstly, we scale the empirical measures $\delta_{x_1} + \ldots + \delta_{x_n}$ by a factor h, secondly we scale all l-nary interactions by a factor h^l, and thirdly we scale the whole generator by $1/h$ (which effectively means the scaling of time as in the theory of superprocesses, see e.g. [14]). Of course, the second and the third scaling can be combined to yield the scaling of l-nary interactions by h^{l-1}. Moreover, allowing for an arbitrary mean-field interaction means to let the transition probabilities P to depend on the measure (on the whole field). After this scaling and the inclusion of an additional mean field interaction, operator (3.3) takes the form

$$\begin{aligned} \Lambda^h f(h\nu) = \frac{1}{h} \sum_{l=0}^{k} h^l \sum_{I \subset \{1, \ldots, n\}, |I| = l} \sum_{m=0}^{\infty} \int \Big[f\Big(h\nu - \sum_{i \in I} h\delta_{x_i} \\ + h\delta_{y_1} + \ldots + h\delta_{y_m} \Big) - f(h\nu) \Big] P(h\nu; \mathbf{x}_I; dy_1 \ldots dy_m) \end{aligned} \tag{3.4}$$

(we have omitted the index k from the notation of Λ for brevity) and acts on the space $B(\mathcal{M}^+_{h\delta}(X))$ of functions defined on the set $\mathcal{M}^+_{h\delta}(X)$ of measures of the form $h\nu = h\delta_{x_1} + \ldots + h\delta_{x_n}$. This generator defines our basic Markov model of exchangeable particles with (h-scaled) k-nary mean-field interaction of pure jump type.

3.2. Derivation of kinetic equations for pure jump models

By Proposition 2.1, operator (3.4) can be written in the form

$$\frac{1}{h} \sum_{l=0}^{k} \frac{1}{l!} \sum_{m=0}^{\infty} \int [f(h\nu - h\delta_{z_1} - \ldots - h\delta_{z_l} + h\delta_{y_1} + \ldots + h\delta_{y_m}) - f(h\nu)]$$

$$\times P(h\nu; z_1, \ldots, z_l; dy_1 \ldots dy_m) \prod_{j=1}^{l} (h\delta_{x_1} + \ldots + h\delta_{x_n})(z_j)$$

$$+ \frac{1}{h} \sum_{\Gamma} \alpha_\Gamma h^{l-p} \int [f(h\nu - h\gamma_1\delta_{z_1} - \ldots - h\gamma_p\delta_{z_p} + h\delta_{y_1} + \ldots + h\delta_{y_m}) - f(h\nu)]$$

$$\times P(h\nu; z_1, \ldots, z_1, \ldots, z_p, \ldots, z_p; dy_1 \ldots dy_m) \prod_{j=1}^{p} (h\delta_{x_1} + \ldots + h\delta_{x_n})(z_j), \tag{3.5}$$

where $\sum_\Gamma$ is the sum over all Young schemes $\Gamma = \{1 \le \gamma_1 \le \ldots \le \gamma_p\}$ with $\gamma_p > 1$ and $\gamma_1 + \ldots + \gamma_p = l \le k$, and $P(z_1, \ldots, z_1, \ldots z_p, \ldots z_p; dy_1 \ldots dy_m)$ means that the first γ_1 arguments of P equal z_1, the next γ_2 arguments equal z_2 etc. As $h \to 0$ and $h\delta_{x_1} + \ldots + h\delta_{x_n}$ tends to some finite measure μ (i.e. the number of particles tends to infinity, but the "whole mass" remains finite due to the scaling of each atom), operator (3.5) tends to

$$(\Lambda f)(\mu) = \sum_{l=0}^{k} \frac{1}{l!} \sum_{m=0}^{\infty} \int \Big[-\frac{\delta f}{\delta\mu(z_1)} - \ldots - \frac{\delta f}{\delta\mu(z_l)} + \frac{\delta f}{\delta\mu(y_1)} + \ldots + \frac{\delta f}{\delta\mu(y_m)} \Big]$$

$$\times P(\mu; z_1, \ldots, z_l; dy_1 \ldots dy_m) \prod_{j=1}^{l} \mu(dz_j). \tag{3.6}$$

The equation $\dot f = \Lambda f$ describing the evolution of functions on $\mathcal{M}^+(X)$ is a first order partial differential equation in functional derivatives. The kinetic equations we are looking for are the characteristics of this equation describing the evolution of measures themselves. Formally these characteristics can be found by writing down the corresponding evolution of a measure-valued function $f(\mu) = \mu$. As formally

$$\frac{\delta\mu}{\delta\mu(z)} = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(\mu + \varepsilon\delta_z - \mu) = \delta_z,$$

the evolution of measures is given by the equation

$$\frac{d}{dt}\mu_t(dz) = \sum_{l=0}^{k} \frac{1}{l!} \sum_{m=0}^{\infty} \int [-\delta_{z_1}(z) - \ldots - \delta_{z_l}(z) + \delta_{y_1}(z) + \delta_{y_m}(z)]$$

$$\times P(\mu; z_1, \ldots, z_l; dy_1 \ldots dy_m) \prod_{j=1}^{l} \mu_t(dz_j)$$

(where it would be better for consistency to write $\delta_y(dz)$ instead of $\delta_y(z)$, but we kept here the standard convention denoting the Dirac measure as a generalized function). Using notations (2.4) and (2.6), one can rewrite this equation in the following concise form

$$\frac{d}{dt}\mu_t(\cdot) = \int (\mathbf{1}^+_{(\cdot)}(\mathbf{y}) - \mathbf{1}^+_{(\cdot)}(\mathbf{z}))\, P(\mu;\mathbf{z};d\mathbf{y})\mu_t^{\tilde{\otimes}}(d\mathbf{z}) \tag{3.7}$$

(where $(\cdot)$ denotes an arbitrary Borel set), which is precisely our general kinetic equation in the strong form.

It is even simpler to obtain from (3.6) the weak form of this equation. As for a linear function

$$f_g(\mu) = \int g(x)\mu(dx) \tag{3.8}$$

on $\mathcal{M}(X)$ one has $\delta f_g/\delta\mu(z) = g(z)$ and hence

$$(\Lambda f_g)(\mu) = \sum_{l=0}^{k} \frac{1}{l!} \int [g^+(\mathbf{y}) - g(z_1) - \ldots - g(z_l)]\, P(\mu; z_1, \ldots, z_l; d\mathbf{y}) \prod_{j=1}^{l} \mu(dz_j) \tag{3.9}$$

(where we used notation (2.6)), the equation $\dot{f}_\mu = \Lambda f_\mu$ becomes

$$\frac{d}{dt}\int g(z)\mu_t(dz) = \int (g^+(\mathbf{y}) - g^+(\mathbf{z}))P(\mu;\mathbf{z};d\mathbf{y})\mu_t^{\tilde{\otimes}}(d\mathbf{z}). \tag{3.10}$$

Of course, (3.7) is just (3.10) with g being an indicator function. We say that $\mu_t(dx)$ is a *weak solution* of the kinetic equation (3.7) if (3.10) holds for any $g(z) \in C_c(X)$, where the integral is well defined in the Lebesgue sense.

3.3. General limiting equations

We worked first through pure jump models to show the main idea in a simpler situation, and also because our rigorous convergence results in the next section will be devoted only to this case leaving other models for further work. However, on the level of a formal limit, writing down general equations does not involve serious complications. These equations appear if one allows for not necessarily jump interactions for particles. As shown in [4] under rather general assumptions, the parts of the generators of Feller processes on $\mathcal{X}$ that are not of pure jump type generate the processes that preserve the number of particles. Hence the general Feller process on $\mathcal{X}$ of k-nary interacting particles can be specified by a transition kernel P from $S\mathcal{X}$ to $S\mathcal{X}$ as above and a collection $A = (A^1, A^2, \ldots, A^k)$ of the generators of (symmetric) Feller processes on $X, X^2, \ldots, X^k$ respectively, and is then defined by the generator

$$(\tilde{G}_{\leq k} f)(x_1, \ldots, x_n) = (G_{\leq k} f)(x_1, \ldots, x_n) + \sum_{I \subset \{1,\ldots,n\}} (A_I f)(x_1, \ldots, x_n),$$

where $G_{\leq k}f$ is given by (3.3) and A_I is the operator A^l with $l=|I|$ that acts on the variables $\mathbf{x}_I$. Instead of (3.4) one gets now a more general expression

$$\tilde{\Lambda}^h f(h\nu) = \frac{1}{h}\sum_{l=0}^k h^l \sum_{I\subset\{1,\dots,n\},|I|=l} \Bigg((A_I(h\nu))f(h\nu) + \sum_{m=0}^{\infty}\int \Big[f\Big(h\nu - \sum_{i\in I} h\delta_{x_i} + h\delta_{y_1}+\dots+h\delta_{y_m}\Big) - f(h\nu)\Big] P(h\nu;\mathbf{x}_I;dy_1\dots dy_m)\Bigg), \tag{3.11}$$

where A can now also depend on the field $h\nu$. One can work out how A acts on the functionals on measures, but for the derivation of limiting equation one only needs to know how it acts on linear functionals (3.8). Namely, one sees from Proposition 2.1 that for $f_g(h\nu)$ with $h\nu = h\delta_{x_1}+\dots+h\delta_{x_n}$

$$\frac{1}{h}h^l \sum_{I\subset\{1,\dots,n\},|I|=l} (A_I f_g)(h\nu) = \frac{1}{l!}\int (Ag^+)(z_1,\dots,z_l)\prod_{j=1}^l (h\nu)(dz_j) + O(h),$$

where an obvious notation $(Ag^+)(z_1,\dots,z_l) = (A^l g^+)(z_1,\dots,z_l)$ is used, and consequently (3.10) generalizes to the equation

$$\frac{d}{dt}\int g(z)\mu_t(dz) = \int\Big[(A(\mu)g^+)(\mathbf{z}) + \int (g^+(\mathbf{y}) - g^+(\mathbf{z}))\,P(\mu;\mathbf{z};d\mathbf{y})\Big]\mu_t^{\tilde{\otimes}}(d\mathbf{z}) \tag{3.12}$$

which represents a weak form of the limiting equation in its most general form.

3.4. Various forms of kinetic equations

Due to the symmetry, equation (3.10) can be equivalently written as

$$\frac{d}{dt}\int g(z)\mu_t(dz) = \sum_{l=0}^k\sum_{m=1}^{\infty}\int_{z_1,\dots,z_l,y_1,\dots,y_m} \frac{1}{l!}(mg(y_1) - lg(z_1)) \times \mu_t(dz_1)\dots\mu_t(dz_l)\,P(\mu_t;z_1,\dots,z_l;dy_1\dots dy_m). \tag{3.13}$$

This allows to write the following equivalent form of the strong equation (3.7):

$$\frac{d}{dt}\mu_t(dz) = -\sum_{l=1}^k \frac{1}{(l-1)!}\mu_t(dz)\int_{z_1,\dots,z_l} \mu_t(dz_1)\dots \times \mu_t(dz_{l-1})\,P(\mu;z,z_1,\dots,z_{l-1})$$

$$+\sum_{l=0}^{k}\frac{1}{l!}\sum_{m=1}^{\infty}m\int_{z_1,\dots,z_l,y_1,\dots,y_{m-1}}\mu_t(dz_1)\dots$$
$$\times\,\mu_t(dz_l)\,P(\mu;z_1,\dots,z_l;dzdy_1\dots dy_{m-1}). \qquad (3.14)$$

As one is often interested in measures having densities with respect to some given reference measure dx on X (say, Lebesgue measure on $\mathbf{R}^n$), it is useful to have an equation on the densities. If, say, there is no additional mean field interaction (e.g. P does not depend on μ), these equations can be obtained whenever the r.h.s. $\pi(\mu_t)$ of (3.14) preserves the subspace of measures having densities with respect to dx, for example, if there exist measures $Q(y;dz_1\dots dz_l)$ such that

$$\sum_{m=1}^{\infty}m\,dz_1\dots dz_l\int_{y_1,\dots,y_m}P(z_1,\dots,z_l;dydy_1\dots dy_m)=Q(y;dz_1\dots dz_l)dy \qquad (3.15)$$

for all l (a criterion for the existence of such Q is given in [4]). Then clearly equation (3.14) holds for measures of the form $\mu_t(dz)=\phi_t(z)dz$ whenever the densities ϕ_t satisfy the following kinetic equation on densities

$$\dot{\phi}_t(z)=-\sum_{l=1}^{k}\frac{1}{(l-1)!}\phi_t(z)\int_{z_1,\dots,z_{l-1}}\phi_t(z_1)\dots$$
$$\times\,\phi_t(z_{l-1})\,P(z,z_1,\dots,z_{l-1})dz_1\dots dz_{l-1}$$
$$+\sum_{l=0}^{k}\frac{1}{l!}\int_{z_1,\dots,z_l}\phi_t(z_1)\dots\phi_t(z_l)Q(z;dz_1\dots dz_l). \qquad (3.16)$$

In discrete case, when $X=\mathbf{Z}^+=\{1,2,\dots\}$, say, and if there is no mean field interaction, (3.14) takes the form

$$\dot{\nu}_j=-\sum_{l=1}^{k}\frac{1}{(l-1)!}\sum_{m=0}^{\infty}\sum_{i_1,\dots,i_{l-1}}\sum_{j_1,\dots,j_m}P^{j_1,\dots,j_m}_{j,i_1,\dots,i_{l-1}}\,\nu_j\nu_{i_1}\dots\nu_{i_l}$$
$$+\sum_{l=0}^{k}\frac{1}{l!}\sum_{m=1}^{\infty}m\sum_{i_1,\dots,i_l}\sum_{j_1,\dots,j_{m-1}}P^{j,j_1,\dots,j_{m-1}}_{i_1,\dots,i_l}\,\nu_{i_1}\dots\nu_{i_l}, \qquad (3.17)$$

where we denoted by $\nu=\{\nu_j(t)\}$ the unknown measure on $\mathbf{Z}^+$ and omitted the dependence on t for brevity. By shifting the indexes, this can be written in the concise form as

$$\dot{\nu}_j=\sum_{l=0}^{k}\frac{1}{l!}\sum_{m=0}^{\infty}\sum_{i_1,\dots,i_l}\sum_{j_1,\dots,j_m}\nu_{i_1}\dots\nu_{i_l}[(m+1)P^{j,j_1,\dots,j_m}_{i_1,\dots,i_l}-P^{j_1,\dots,j_m}_{j,i_1,\dots,i_l}\nu_j], \qquad (3.18)$$

3.5. Examples without an explicit mean field interaction

The most important particular case of discrete equations (3.18) corresponds to the mass preserving (or mass exchange) processes, where the indexes in (3.18) are interpreted as masses and it is supposed that the rates $P^{j_1,\ldots,j_m}_{i_1,\ldots,i_l}$ may not vanish only if $i_1+\ldots+i_l = j_1+\ldots+j_m$. The mathematical theory of the equations of these kind was developed in [21]. These equations correspond to a variety of processes including classical Smoluchovski's coagulation fragmentation equations with discrete mass distribution and its generalization with not necessary binary coagulations, coagulation equations with collision breakage, some known models of evolution of raindrops size spectra etc., see a fuller discussion and the relevant bibliography in [21]. Other types of examples come from biology. For instance, the classical Lotka–Volterra predator-prey model (and its various generalizations) presents a simple particular case of (3.18).

As a basic example of general pure jump interactions, let us distinguish the processes that combine pure coagulations of no more than k particles, spontaneous fragmentation in no more than k pieces, and collisions (or collision breakages) of no more than k particles (with an arbitrary k). These processes are specified by the transition kernels $P^l_1(z_1,\ldots,z_l) = K_l(z_1,\ldots,z_l;dy)$, $l=2,\ldots,k$, called the coagulation kernels, the transition kernels

$$P^1_m(z;dy_1\ldots dy_m) = F_m(z;dy_1\ldots dy_m), \quad m=2,\ldots,k,$$

called the fragmentation kernels, and the kernels

$$P^l_l(z_1,\ldots,z_l;dy_1\ldots dy_l) = C_l(z_1,\ldots,z_l;dy_1\ldots dy_2), \quad l=2,\ldots,k,$$

called the collision kernels, all other P^m_l are supposed to vanish. Then equation (3.10) takes the form

$$\begin{aligned}
&\frac{d}{dt}\int g(z)\mu_t(dz)\\
&\quad= \sum_{l=2}^k \frac{1}{l!}\int\limits_{z_1,\ldots,z_l,y} [g(y)-g(z_1)-\ldots-g(z_m)]\,K_l(z_1,\ldots,z_l;dy)\prod_{j=1}^l \mu_t(dz_j)\\
&\quad+\sum_{m=2}^k \int\limits_{z,y_1,\ldots,y_m} [g(y_1)+\ldots+g(y_m)-g(z)]\,F_m(z;dy_1\ldots dy_m)\,\mu_t(dz)\\
&\quad+\sum_{l=2}^k \int [g(y_1)+\ldots+g(y_l)-g(z_1)-\ldots\\
&\quad-g(z_l)]C_l(z_1,\ldots,z_l;dy_1\ldots dy_l)\prod_{j=1}^l \mu_t(dz_j). \qquad (3.19)
\end{aligned}$$

In usual mechanical models, some function L on X is given that is preserved by the kernels K, F, C, which means that, say, the measures $K_l(z_1, \ldots, z_l; \cdot)$ are supported on the set $\{y : L(y) = L(z_1) + \ldots + L(z_l)\}$. In particular, in case of pure mass exchange processes, a particle is characterized by its mass only, $X = \mathbf{R}^+$ and $L(x) = x$ is interpreted as the mass of a particle. On the other hand, to get a (spatially homogeneous) Boltzmann equation, one sets all K_l and F_l to be zero, $X = \mathbf{R}^3$ (the space of velocities of particles) and takes the function L to be the kinetic energy v^2 of a particle with velocity v. The resulting equation can be called k-nary Boltzmann equation (see [23] for its detailed discussion).

Classical Smoluchovski's coagulation equation corresponds to the case when only $P_1^2 = K_2$ in (3.19) does not vanish. The mathematical theory of such equation (existence and uniqueness of solutions, convergence of stochastic approximations) was developed recently in [30, 31]. The existence and uniqueness of solutions for the case with vanishing C, $X = \mathbf{R}^+$ and K_l and F_l being non-vanishing only for $l = 2$, i.e. for the classical coagulation-fragmentation model

$$\frac{d}{dt} \int g(z)\mu_t(dz) = \int_{z_1, z_2} [g(z_1 + z_2) - g(z_1) - g(z_2)] K(z_1, z_2)\mu_t(dz_1)\mu_t(dz_2)$$
$$+ \int_0^z dy \int_0^\infty \mu_t(dz)[g(y) + g(z - y) - g(z)]F(z, y) \qquad (3.20)$$

was obtained in [13] for solution measures μ_t having densities with respect to Lebesgue measure and also finite exponential moments. Though in physical literature the model with collision breakage is well known (see e.g. [10,12,25,34]), mathematical results on even simplest pure binary collision breakage (only C_2 does not vanish in (3.16)) are obtained for simplified models of discrete mass distribution in [29], see also [21] and [26]. The full situation with non-vanishing F, K, C was not seemingly investigated mathematically even for pure binary case. As we noted in our general setting (when X is not necessarily $\mathbf{R}_+$ and L is not necessarily the mass) this model comprises also the Boltzmann collisions.

As one sees by inspection, the more general equation (3.12) (even without explicit dependence of A and P on μ) contains already all classical kinetic equations of statistical mechanics, namely that of Landau and Vlasov, and the spatially nontrivial equations of Boltzmann and Smoluchovski. A mollified k-nary versions of the Boltzmann and Smoluchovski equations were analyzed in [23].

Kinetic equations (3.12) without a mean field interaction (i.e. without explicit dependence of A and P on μ), were first obtained in [4] (in a slightly different form) by a different method. In fact, we developed two such methods (both being formal, e.g. without any rigorous convergence or existence results), one was suggested in [3] (where the case of bounded generators of quantum semigroups was considered) and was based on the study of the evolution of the generating functionals and another was based on the idea of the propagation of chaos (see e.g. [5] or [35]). The method used in the present paper is more appropriate for proving rigorous theorems, as we shall show further for the case of equation (3.10).

3.6. Examples with a mean field interaction included. Replicator dynamics

Including mean field interaction allows for essentially more general models. In this way one can get various kinds of nonlinearity, for example those appearing in hydrodynamic limits of exclusion processes, see e.g. [16, 27], and references therein. What is more important for our present purpose, is the possibility to include the general replicator dynamics (and other evolutionary dynamics) from evolutionary game theory, as we shall explain now.

Recall that in order to make the influences of l-nary interactions with various l of the same order, we have to re-scale the l-nary part of the generator by h^{l-1}. Another reasonable scaling (normally used in evolutionary games) is by normalizing the higher order interactions by the number of particles, i.e. by division l-nary part of (3.3) by $n^{l-1} = (\|h\nu\|/h)^{l-1}$. This yields

$$\Lambda^h f(h\nu) = \frac{1}{h}\sum_{l=0}^{k} h^l \sum_{I\subset\{1,\dots,n\},|I|=l} \sum_{m=0}^{\infty} \int \Big[f\Big(h\nu - \sum_{i\in I} h\delta_{x_i} + h\delta_{y_1} + \dots + h\delta_{y_m}\Big) - f(h\nu)\Big] \frac{P(h\nu;\mathbf{x}_I;dy_1\dots dy_m)}{\|h\nu\|^{(l-1)}} \tag{3.21}$$

instead of (3.4), and hence the limiting equation

$$\frac{d}{dt}\int g(z)\mu_t(dz) = \int (g^+(\mathbf{y}) - g^+(\mathbf{z}))P(\mu;\mathbf{z};d\mathbf{y})\Big(\frac{\mu_t}{\|\mu_t\|}\Big)^{\tilde\otimes}(d\mathbf{z})\|\mu_t\| \tag{3.22}$$

for pure jump processes and to a similar modification of (3.12) in general case. Of course, (3.22) can be written again in form (3.10), so it does not give anything really new, it just specifies certain type of a mean-field interaction. The point is that the standard replicator dynamics of evolutionary game theory describing binary interacting (playing two-persons game) species in a population with finite number of strategies is written precisely in form (3.22), see e.g. [9,17,36]. There also appeared recently equations of these type describing coagulation processes in biological context (Phytoplankton dynamics), see [1]. Generalization to k-nary interactions (species playing a k person game), arbitrary interactions laws and state spaces given by equation (3.22) seems to be quite relevant in the biological context, see [33], it also allows to analyze animals living in groups or in large families, see [20] for further discussion. It is worth noting that if one considers interactions of a fixed order (i.e. (3.21) includes a term with only one index l), then re-scaling of interactions leading to (3.22) is equivalent to a time change in (3.10) (a performance of this reduction in evolutionary biology is the well known trajectory-wise equivalence of Lotka – Volterra model and replicator dynamics, see [17]). But if the interactions of various order are present, such a simple reduction does not seem to be possible anymore.

It seems of interest to exploit the above (biologically tested) scaling on the usual models of stochastic mechanics. Say for binary coagulation-fragmentation model described by equation (3.20), this scaling would lead to the equation

$$\frac{d}{dt}\int g(z)\mu_t(dz) = \int\limits_{z_1,z_2} \|\mu_t\|^{-1}[g(z_1+z_2)-g(z_1)-g(z_2)] \times K(z_1,z_2)\mu_t(dz_1)\mu_t(dz_2) + \int\limits_0^z dy \int\limits_0^\infty \mu_t(dz)[g(y)+g(z-y)-g(z)]F(z,y), \qquad (3.23)$$

which has a r.h.s. of linear growth (and not quadratic as for usual Smoluchovski equation). It would be interesting to test this new model by real chemical experiments. It is worth being noted that though linearity of growth of the r.h.s. in (3.22) seems to be a simplification, it causes other technical difficulties arising from the possibility of small $\|\mu_t\|$. Hence in our future analysis we often have to consider (3.10) and (3.22) separately.

4. Existence of solutions and convergence of stochastic approximations

This section is devoted to a justification of heuristic arguments of the previous section for the case of particle systems with pure jump interaction. Under some additional assumptions we are going to prove the existence of global solutions to (3.10) and the convergence of approximating exchangeable particle systems to these solutions. The following definitions will play the crucial role in our analysis.

Let L be a non-negative function on X. The number $L(x)$ will be called the size of a particle x (L stands for the mass in mass exchange models like coagulation-fragmentation, and L stands for the kinetic energy when modelling Boltzmann type collisions). We say that the transition kernel $P = P(\mu;\mathbf{x};d\mathbf{y})$ in (3.4) or (3.6) is *L-subcritical* (respectively *L-critical*), if

$$\int (L^+(\mathbf{y}) - L^+(\mathbf{x}))P(\mu;\mathbf{x};d\mathbf{y}) \le 0 \qquad (4.1)$$

for all $\mathbf{x}$ and μ (respectively if the equality holds). We say that $P(\mu;\mathbf{x};d\mathbf{y})$ in (3.4) is *L-preserving* (respectively *L-non-increasing*) if the measure $P(\mu;\mathbf{x};d\mathbf{y})$ is supported on the set $\{\mathbf{y} : L^+(\mathbf{y}) = L^+(\mathbf{x})\}$ (respectively $\{\mathbf{y} : L^+(\mathbf{y}) \leq L^+(\mathbf{x})\}$). We say that $P(\mu; \mathbf{x}; d\mathbf{y})$ is *l-nary L-subcritical* (L-critical, L-preserving, L-non-increasing), if the corresponding property holds only for $\mathbf{x} \in X^l$. Clearly, if $P(\mu; \mathbf{x}; d\mathbf{y})$ is L-preserving (respectively L-non-increasing), then it is also L-critical (respectively L-subcritical).

We shall say that our transition kernel P is *multiplicatively L-bounded* or *$L^{\otimes}$-bounded* (respectively *additively L-bounded* or *L^+-bounded*) whenever $P(\mu; \mathbf{x}) \leq cL^{\otimes}(\mathbf{x})$ (respectively $P(\mu; \mathbf{x}) \leq cL^+(\mathbf{x})$) for all μ and $\mathbf{x}$ and some constant $c > 0$, where we used the notation (2.5), (2.6), (3.2). In the future, we shall always take $c = 1$ here for brevity. We shall say that P is *strongly multiplicatively L-bounded* whenever $P(\mu; \mathbf{x}) = o(1)_{\mathbf{x}\to\infty} L^{\otimes}(\mathbf{x})$, i.e. if

$$P(\mu; x_1, \ldots, x_l) = o(1)_{(x_1,\ldots,x_l)\to\infty} \prod_{j=1}^{l} L(x_j) \tag{4.2}$$

for $l = 1, \ldots, k$. We shall need also some conditions that forbid the creation of dust, i.e. of a large number of small particles. Various conditions of this kind are possible. We choose the following no dust condition (for a given L):

> (ND) the number of particles that can be created by a single act of interaction is uniformly bounded by some number m_0, and P is l-nary $\mathbf{1}$-subcritical for $l \geq 2$.

Remark 4.1. We choose the simplest possible (ND) condition. Various modifications are possible. For instance, one can use P that are l-nary $\mathbf{1}_{L(\cdot)<r}$-subcritical for all $l \geq 2$ instead of being l-nary $\mathbf{1}$-subcritical.

Remark 4.2. The notion of L-subcriticality is borrowed from the theory of branching processes and superprocesses, where it is used usually with $L = \mathbf{1}$ (see also [20] for interacting particles). The assumptions of $L^{\otimes}$- and L^+-boundedness constitute natural generalizations of two basic assumptions on the coagulation kernel used in the theory of Smoluchovski's equation (see e.g. [2] and [30]). Condition (ND) is trivially satisfied for our basic model (3.19).

Remark 4.3. Clearly $(1+L^{\alpha})^+$-boundedness implies $(1+L)^{\otimes}$-boundedness whenever $\alpha \in [0,1]$, and it implies strong $(1+L)^{\otimes}$-boundedness whenever $\alpha < 1$.

We begin now our analysis of the Markov processes defined by (3.4) and (3.21).

Proposition 4.1. *Suppose the transition kernel $P(\mu; \mathbf{x}, \cdot)$ is a continuous function from $\mathcal{M}^+(X) \times S\mathcal{X}$ to $\mathcal{M}^+(S\mathcal{X})$ (where $\mathcal{M}^+(S\mathcal{X})$ is considered with its weak topology), is L-subcritical for some continuous non-negative function L on X such that every set $\{x : L(x) \leq a\}$ is a compact for any a (for example,*

$L(x) \to \infty$ whenever $x \to \infty$), is $(1+L)^{\otimes}$-bounded and condition (ND) *holds. Suppose also for simplicity that $P^0(dy) = 0$, i.e. that no spontaneous input in the system is possible. Then*

(i) *the Markov processes $Z^{h\nu}(t)$ ($h\nu$ denotes as usual the starting point) in $\mathcal{M}^+_{h\delta}(X) \subset \mathcal{M}^+(X)$ are correctly and uniquely defined by generators either of type* (3.4) *or* (3.21)*, and the processes $f_L(Z^{h\nu}(t))$ are non-negative supermartingales, where f_L is defined by* (3.8)*;*

(ii) *given arbitrary $b > 0$, $T > 0$, for a family $Z^{h\nu}(t)$ with $h \in (0,1]$ and initial $h\nu$ with $h\int(\mathbf{1}+L)(x)\nu(dx) \le b$*

$$\mathsf{P}\big(\sup_{t\in[0,T]} f_{L+1}(Z^{h\nu}(t)) > r\big) \le \frac{C(T,b,m_0)}{r} \tag{4.3}$$

for all $r > 0$, where the constant $C(T,b,m_0)$ does not depend on h; moreover, such a family $Z^{h\nu}(t)$ enjoys the compact containment condition, i.e. for arbitrary $\eta > 0$, $T > 0$ there exists a compact subset $\Gamma_{\eta,T} \subset \mathcal{M}^+(X)$ for which

$$\inf_{h\nu} \mathsf{P}\big(Z^{h\nu}(t) \in \Gamma_{\eta,T} \text{ for } 0 \le t \le T\big) \ge 1-\eta; \tag{4.4}$$

(iii) *if the family of measures $\nu = \nu(h)$ is such that $f_{1+L}(h\nu)$ is uniformly bounded, then the family of processes $Z^{h\nu}(t)$ defined by* (3.4) *is tight as a family of processes with sample paths in $D_{\mathcal{M}^+(X)}[0,\infty)$; the same holds for processes $Z^{h\nu}(t)$ defined by* (3.21) *whenever P is $(1+L)^+$- bounded.*

Proof. (i) By L-subcriticallity and by (3.4), (4.1)

$$\Lambda^h f_L(h\nu) \le 0, \quad \nu \in \mathcal{M}^+_\delta(X).$$

By **1**-subcriticallity of the interactions of order $l \ge 2$

$$\Lambda^h f_{\mathbf{1}}(h\nu) \le h\sum_{i=1}^n \int (\mathbf{1}^+(\mathbf{y}) - 1)P(h\nu; x_i; d\mathbf{y}),$$

for $\nu = \delta_{x_1} + \ldots + \delta_{x_n}$ and hence

$$\Lambda^h f_{\mathbf{1}}(h\nu) \le m_0 \sum_{i=1}^n (\mathbf{1}+L)(x_i) = m_0 f_{\mathbf{1}+L}(h\nu),$$

as the number of particles created in one go is bounded by m_0. Moreover, the intensity q of the q-pair corresponding to the generator (3.4) equals

$$q(h\nu) = \frac{1}{h}\sum_{l=1}^k h^l \sum_{I\subset\{x_1,\ldots,x_n\},|I|=l} \int P(h\nu; \mathbf{x}_I; d\mathbf{y}) \le \frac{1}{h}\sum_{l=1}^k \frac{1}{l!}(f_{1+L}(h\nu))^l \tag{4.5}$$

for $\nu = \delta_{x_1} + \ldots + \delta_{x_n}$. In case of (3.21) the same estimate holds with an additional multiplier $h^{-(l-1)}$. Hence the conditions of Proposition 2.2 (ii) hold with $f_{\mathbf{1}+L}$ playing the role of the barrier ψ.

(ii) Estimate (4.3) follows from (i) and formula (2.11) of Proposition 2.1. This implies that $f_{1+L}(Z^{h\nu}(t))$ is uniformly bounded with a probability arbitrary closed to one. This implies (4.4) by the Prokhorov criterion of compactness for a family of measures and because $L(x) \to \infty$, as $x \to \infty$.

(iii) By the Jacubovski criterion (see e.g. [15] or [11]), when the compact containment condition (4.4) is proved, in order to get tightness it is enough to show the tightness of the family of the real valued processes $f(Z^{h\nu}(t))$ (as a family of processes with sample paths in $D_{\mathbf{R}}[0,\infty)$) for any f from a dense subset (in the topology of uniform convergence on compact sets) of $C_b(\mathcal{M}^+(X))$. By Weierstrass theorem, it is thus enough to verify tightness of $f(Z^{h\nu}(t))$ for f from the algebra generated by f_g of type (3.9) with $g \in C_c(X)$. Let us show it for $f = f_g$ (as seen from the proof, it is straightforward to generalize it to the sums and products of these functions). By Propositions 2.2 and 4.1, $f_g(Z^{h\nu}(t))$ is a supermartingale and moreover (by (2.9)), the process

$$M_g(t) = f_g(Z^{h\nu}(t)) - f_g(h\nu) - \int_0^t \Lambda^h f_g(Z^{h\nu}(s))\,ds \tag{4.6}$$

is a martingale for any $g \in C_c(X)$. By standard criteria of tightness of real valued supermartingales, to get a tightness for $f_g(Z^{h\nu}(t))$, one needs to estimate the oscillations of the predictable finite variation component of $f_g(Z^{h\nu}(t))$, which is given by the integral in (4.6) and for which such an estimate is thus obvious, and for the quadratic variation $[M_g(t)]$ of the martingale (4.6). In fact, we shall show that for an arbitrary ε,

$$\mathsf{P}\big([M_g(t)] - [M_g(s)] \le \sigma h(t-s)\big) \ge 1 - \varepsilon \tag{4.7}$$

with some constant σ uniformly for all $0 \le s \le t \le T$ and an arbitrary T, and this implies the required tightness.

Since the process $Z^{h\nu}(t)$ is a pure jump process,

$$[M_g(t)] = \sum_{s \le t} \big(\Delta f_g(Z^{h\nu}(s))\big)^2, \tag{4.8}$$

where $\Delta Z(s) = Z(s) - Z(s_-)$ denotes the jump of a process $Z(s)$. Since the number of particles created in one go is bounded by m_0, it follows from (3.8) that

$$|\Delta f_g(Z^{h\mu}(s))|^2 \le \|g\|^2 (m_0 + k)^2 h^2$$

for any s. Clearly now in order to get (4.7) it is enough to show that the number of jumps on any interval $[s,t]$ is of order $(t-s)/h$ uniformly for all $s \le t \le T$

with any T with probability arbitrary close to one. But as our Markov process is a pure jump process, the number of jumps on the interval $[s,t]$ is controlled by the product of $(t-s)$ and the maximal intensity on $[0,T]$ (for example, by the standard Levy formula for (4.8), see e.g. [6]). So, it is enough to prove that the maximal intensity is of order $1/h$. But for $\nu = \delta_{x_1} + \ldots + \delta_{x_n}$, one gets from (4.5) that in case of generator (3.4) $q(h\nu)$ is of order $1/h$ with the probability arbitrary close to one due to the compact containment condition. This proves (4.6) and hence completes the proof of Proposition 4.1. In case of generator (3.21) and assuming $(1+L)^+$-boundedness of P one gets instead of (4.5) the estimate

$$q(h\nu) \leq \frac{1}{h} \sum_{l=1}^{k} \frac{1}{l!} (f_{1+L}(h\nu)),$$

which again implies that $q(h\nu)$ is of order h^{-1} and completes the proof in case of generator (3.21). □

We can now prove our first existence result for equation (3.10).

Theorem 4.1. *Under assumptions of Proposition 4.1 suppose additionally that the family of initial measures $h\nu$ converges (weakly) to a measure $\mu \in \mathcal{M}^+(X)$ and that the kernel P is strongly $(1+L)^{\otimes}$-bounded. Then there is a subsequence of the family of processes $Z^{h\nu}(t)$ defined by generator (3.4) that weakly converges to a global non-negative solution μ_t of the integral version of the weak equation (3.10), i.e.*

$$\int g(z)\mu_t(dz) - \int g(z)\mu_0(dz) - \int_0^t ds \int (g^+(\mathbf{y}) - g^+(\mathbf{z}))P(\mathbf{z}; d\mathbf{y})\mu_s^{\tilde{\otimes}}(d\mathbf{z}) = 0 \tag{4.9}$$

for all t and all $g \in C_c(X)$.

Proof. By Proposition 4.1 we can choose a sequence of positive numbers tending to zero such that the family $Z^{h\nu}(s)$ is weakly converging as $h \to 0$ and belong to this sequence. Let us denote the limit by μ_t and prove that it satisfies (4.9). By Skorohod's theorem, we can and will assume that $Z^{h\nu}(s)$ converges to μ_t almost surely. The idea now is to pass to the limit in equation (4.6). Due to (4.7), the martingale on the left hand side of (4.6) tends to zero almost surely. Moreover, due to our estimates of (4.8), any limiting process μ_t has continuous sample paths almost surely. The positivity of μ_t is obvious as it is a limit of positive measures.

Clearly, the first two terms on the r.h.s. of (4.6) tend to the first two terms on the l.h.s. of (4.9). So, we need to show that the integral on the r.h.s. of (4.6) tends to the last integral on the l.h.s. of (4.9). Let us show first that almost surely

$$|\Lambda^h f_g(Z^{h\nu}(s)) - \Lambda f_g(Z^{h\nu}(s))| \to 0, \quad h \to 0, \tag{4.10}$$

uniformly for $s \in t$. As by (3.5) for any $\eta = \delta_{v_1} + \ldots + \delta_{v_m}$,

$$\begin{aligned}
\Lambda^h f_g(h\eta) &= \sum_{l=1}^k h^l \sum_{I \subset \{1,\ldots,m\}, |I|=l} \int \Big(g^+(\mathbf{y}) - \sum_{i \in I} g(v_i) \Big) P(h\eta; \mathbf{v}_I; d\mathbf{y}) \\
&= \sum_{l=0}^k \frac{1}{l!} \int [g^+(\mathbf{y}) - g(z_1) - \ldots - g(z_l)] \\
&\quad \times P(h\eta; z_1, \ldots, z_l; d\mathbf{y}) \prod_{j=1}^l (h\eta(dz_j)) \\
&\quad + \sum_\Gamma \alpha_\Gamma h^{l-p} \int (g^+(\mathbf{y}) - \gamma_1 g(z_1) - \ldots - \gamma_p g(z_p)) \\
&\quad \times P(h\eta; z_1, \ldots, z_1, \ldots, z_p, \ldots, z_p; d\mathbf{y}) \prod_{j=1}^p (h\eta(dz_j)),
\end{aligned}$$

in order to prove (4.10), one needs to show that all

$$h^{l-p} \int_{z_1,\ldots,z_p,\mathbf{y}} P(h\eta; z_1, \ldots, z_1, \ldots, z_p, \ldots, z_p; d\mathbf{y}) \prod_{j=1}^p (h\eta(dz_j)) \tag{4.11}$$

tend to zero as $h \to 0$ uniformly for all $\eta \in \mathcal{M}_\delta^+(X)$ with uniformly bounded $f_{1+L}(h\eta)$.

As the boundedness of $f_L(h\eta)$ implies, in particular, that the support of all η is contained in a set where L is bounded by c/h with some constant c, in order to prove that all terms (4.11) tend to zero it is enough to show that for any i

$$h^i \int_{z: L(z) \leq ch^{-1}} (1 + L(z))^{i+1} \, o(1)_{z \to \infty} (h\eta)(dz) \to 0, \quad h \to 0. \tag{4.12}$$

To this end, let us write the integral in (4.12) as the sum of two integrals over the sets where $L(z) \leq K$ and $L(z) > K$ with some $K > 0$. Then the first term clearly tends to zero, as $L(z)$ is bounded. The second term does not exceed

$$\int_{z: L(z) \geq K} (1 + L(z)) \, o(1)_{K \to \infty} (h\eta)(dz)$$

and tends to zero as $K \to \infty$.

When (4.10) is proved, it remains to show that

$$|\Lambda f_g(Z^{h\nu}(t)) - \Lambda f_g(\mu_t)| \to 0,$$

or more explicitly that the integral

$$\int_0^t ds \int (g^+(\mathbf{y}) - g^+(\mathbf{z}))[P(\mu_s; \mathbf{z}; d\mathbf{y})\mu_s^{\tilde{\otimes}}(d\mathbf{z}) - P(Z^{h\nu}(s); \mathbf{z}; d\mathbf{y})Z^{h\nu}(s)^{\tilde{\otimes}}(d\mathbf{z})] \tag{4.13}$$

tends to zero as $h \to 0$. But from a weak convergence (and the fact that μ_t has continuous sample paths) it follows (see e.g. [15]) that $Z^{Nh}(s)$ converges to μ_s for all $s \in [0, t]$. Hence we need to show that

$$\begin{aligned} &\int P(\mu; \mathbf{z}; d\mathbf{y})[\mu_s^{\tilde{\otimes}}(d\mathbf{z}) - Z^{h\nu}(s)^{\tilde{\otimes}}(d\mathbf{z})] \\ &\quad + \int [P(\mu; \mathbf{z}; d\mathbf{y}) - P(Z^{h\nu}; \mathbf{z}; d\mathbf{y})]Z^{h\nu}(s)^{\tilde{\otimes}}(d\mathbf{z}) \end{aligned} \tag{4.14}$$

tends to zero as $h \to 0$ under condition that $Z^{h\nu}(t)$ weakly converges to μ_t ((4.13) then would follow by the Lebesgue dominated convergence theorem).

Decompose the integrals in (4.14) into the sum of two integrals by decomposing the domain of integration into the domains $\{\mathbf{z} = (z_1, \dots, z_m) : \max L(z_j) \geq K\}$ and its complement. By strong $(1 + L)^{\otimes}$-boundedness, both integrals from (4.14) over the first domain tend to zero as $K \to \infty$. And on the second domain, the integrand is uniformly bounded and hence the weak convergence of $Z^{h\nu}(s)$ to μ_t ensures the smallness of the l.h.s. of (4.14). Theorem 4.1 is proved. □

Theorem 4.2. *Under assumptions of Proposition* 4.1 *suppose additionally that the family of initial measures $h\nu$ converges* (*weakly*) *to a measure $\mu \in \mathcal{M}^+(X)$ and that the kernel P is $(1 + L^\alpha)^+$-bounded with some $\alpha \in [0, 1)$. Then there is a subsequence of the family of processes $Z^{h\nu}(t)$ defined by generator* (3.21) *that weakly converges to a global non-negative solution μ_t of the integral version of the weak equation* (3.22).

Remark 4.4. Note the crucial restriction $\alpha < 1$ here that we are going to dispense with later using deterministic arguments.

Proof. Following the same strategy as above, one has to show (4.10) and to this end the convergence to zero of

$$\begin{aligned} h^{l-p} &\int\limits_{z_1,\dots,z_p,\mathbf{y}} P(h\eta; z_1, \dots, z_1, \dots, z_p, \dots, z_p; d\mathbf{y}) \frac{1}{\|h\eta\|^{l-1}} \prod_{j=1}^p (h\eta(dz_j)) \\ &\leq h^{l-p} \int\limits_{z_1,\dots,z_p,\mathbf{y}} (1 + L^\alpha(z_1)) \frac{1}{\|h\eta\|^{l-1}} \prod_{j=1}^p (h\eta(dz_j)) \\ &= O(h) \, \frac{\int (1 + L^\alpha(z))(h\eta)(dz)}{\|h\eta\|}. \end{aligned}$$

For arbitrary K this can be written as

$$O(h)K + \int_{L(z)\geq K} L^\alpha(z)(h\eta)(dz) = O(h)K + K^{-(1-\alpha)} f_L(h\eta).$$

Choosing first large K and then small h can make this expression arbitrary small. It remains to prove that

$$\int \Big[P(\mu_s; z_1, \ldots, z_l; d\mathbf{y}) \, \frac{\mu_s(dz_1) \ldots \mu_s(dz_l)}{\|\mu_s\|^{l-1}}$$
$$- P(Z^{h\nu}(s); z_1, \ldots, z_l; d\mathbf{y}) \, \frac{Z^{h\nu}(s)(dz_1) \ldots Z^{h\nu}(s)(dz_l)}{\|Z^{h\nu}(s)\|^{l-1}} \Big] \tag{4.15}$$

tends to zero as $h \to 0$ for any l. As in the previous proof we decompose the domain of integration into the domains $\{\mathbf{z} = (z_1, \ldots, z_m) : \max L(z_j) \geq K\}$ and its complement. To deal with the first domain we observe that

$$\int_{z_2, \ldots, z_l, L(z_1) \geq K} L^\alpha(z_1) \Big[\|\mu_s\|^{-(l-1)} \prod_{j=1}^{l} \mu_s(dz_j) + \|Z^{h\nu}(s)\|^{-(l-1)} \prod_{j=1}^{l} Z^{h\nu}(s)(dz_j) \Big]$$
$$\leq K^{-(1-\alpha)} \int L(z_1))[\mu_s(dz_1) + Z^{h\nu}(s)(dz_1)],$$

which tends to zero as $K \to \infty$ (here the assumption $\alpha < 1$ is used). To deal with the second domain in the same way as in the previous theorem, we need to be sure that not only $Z^{h\nu}(s)$ tends to μ_s weakly, but also that $Z^{h\nu}(s)/\|Z^{h\nu}(s)\|$ tends to $\mu_s/\|\mu_s\|$ weakly. To see this let us denote by T the supremum over all t such that μ_t does not vanish. If $T = \infty$, everything becomes obvious. Let $T < \infty$. Then we can complete the proof for all $t < T$. But then μ_t is a solution of the integrated version of equation (3.22). Let us show that it can not vanish at any finite time thus leading to the contradiction with the assumption $T < \infty$ and to the completion of the proof. To see that a solution can not vanish we observe that it is bounded from below by a solution to the equation

$$\frac{d}{dt}\mu_t(dz) = -\sum_{l=1}^{k} \frac{1}{(l-1)!} \frac{\mu_t(dz)}{\|\mu_t\|^{l-1}}$$
$$\times \int_{z_1, \ldots, z_{l-1}} \mu_t(dz_1) \ldots \mu_t(dz_{l-1}) P(\mu; z, z_1, \ldots, z_{l-1})$$

obtained from (3.14) by ignoring the positive part of the r.h.s. In their turn, solutions to this equation are bounded from below by solutions to the equation

$$\frac{d}{dt}\mu_t(dz) = -\sum_{l=1}^{k} \frac{1}{(l-1)!}(1 + L(z))\mu_t(dz)$$

$$-\sum_{l=2}^{k}\frac{1}{(l-2)!}\,\mu_t(dz_1)\frac{\int(1+L(u))\mu(du)}{\|\mu\|}.$$

By Lemma B of the Appendix, the solution to this equation can be found explicitly and it never vanishes. Hence μ_T does not vanish, and consequently our proof is complete. □

The rest of the section will be devoted to the case of L^+-bounded P and solutions having finite L^β-moments. We shall start with the estimates of average moments of L, where in case of non-binary interactions we shall make additional rather restrictive assumptions (ND′) (which will be dispense with later by a deterministic approach; possibly more elaborate probabilistic arguments would allow to get rid of it already at this stage).

Proposition 4.2. *Under the conditions of Proposition* 4.1, *suppose additionally that the kernel P is L-non-increasing and $(1+L)^+$-bounded and in case $k>2$ the following additional condition holds:*

(ND′) *if $k>2$, then either P is $\mathbf{1}$-non-increasing (in particular, no fragmentation is allowed), or $L(x)\geq\varepsilon$ for all x with some $\varepsilon>0$ (i.e. no particle of size less than ε can appear in the system).*

Then for any integer $\beta\geq 1$ (in fact any real $\beta\geq 1$ in case of $k>2$ and condition (ND′))

$$\mathsf{E}(f_{L^\beta}(Z^{h\nu}(s)))\leq a(t)f_{L^\beta}(Z^{h\nu}(0))+b(t) \tag{4.16}$$

uniformly for all $s\in[0,t]$ with an arbitrary t, and with some constants $a(t),b(t)$ depending on t (and β), but not on h.

Proof. As the process is L-non-increasing, the process $Z^{h\nu}(s)$ lives on measures with support on a compact space $\{x\in X: L(x)\leq c/h\}$ with some constant c. Hence $f_{L^\beta}(Z^{h\nu}(s))$ is uniformly bounded (not necessarily uniformly in h) and one can apply Dynkin's formula, i.e. the fact that (4.6) is a martingale for $g=L^\beta$. Hence

$$\mathsf{E}(f_{L^\beta}(Z^{h\nu}(t)))\leq f_{L^\beta}(h\nu)+\int_0^t \mathsf{E}\Lambda^h(f_{L^\beta}(Z^{h\nu}(s)))\,ds. \tag{4.17}$$

Since our kernel P is L-non-increasing,

$$\int(L^\beta)^+(\mathbf{y})P(\mathbf{x},d\mathbf{y})\leq\int(L^+(\mathbf{y}))^\beta P(\mathbf{x},d\mathbf{y})\leq(L^+(\mathbf{x}))^\beta P(\mathbf{x},d\mathbf{y}),$$

and consequently for $\nu = \delta_{x_1} + \ldots + \delta_{x_n}$

$$
\begin{aligned}
(\Lambda^h f_{L^\beta})(h\nu) &\le \frac{1}{h} \sum_{l=1}^k h^l \sum_{|I|=l} [(L^+(\mathbf{x}_I))^\beta - (L^\beta)^+(\mathbf{x}_I)]\, P(\mathbf{x}_I) \\
&\le \sum_{l=1}^k \frac{1}{l!} \int [(L(y_1) + \ldots + L(y_l))^\beta - L^\beta(y_1) - \ldots \\
&\quad - L^\beta(y_l)]\, P(y_1, \ldots, y_l) \prod_{j=1}^l (h\nu(dy_j)),
\end{aligned}
$$

where we used Proposition 2.1. Using the symmetry with respect to permutations of $x_1, \ldots, x_n$, we conclude that $(\Lambda^h f_{L^\beta})(h\nu)$ does not exceed

$$
\begin{aligned}
&\sum_{l=2}^k \frac{1}{(l-1)!} \int [(L(y_1) + \ldots + L(y_l))^\beta - L^\beta(y_1) - \ldots \\
&\quad - L^\beta(y_l)]\, (1 + L(y_1)) \prod_{j=1}^l (h\nu(dy_j)).
\end{aligned}
$$

Using the elementary inequalities

$$
\begin{aligned}
(a+b)^\beta - a^\beta - b^\beta &\le w_\beta^1 (ab^{\beta-1} + ba^{\beta-1}), \\
a((a+b)^\beta - a^\beta - b^\beta) &\le w_\beta^2 (ab^\beta + ba^\beta)
\end{aligned}
$$

(that hold for all $\beta \ge 1$ and positive a, b with some constants w_β^1, w_β^2 depending only on β, see e.g. [7] for a proof of the second one) with $a = L(y_1)$, $b = L(y_2) + \ldots + L(y_l)$, and again the symmetry yields for the last expression the estimate

$$
\sum_{l=2}^k \kappa_\beta^l (L^\beta(y_1) + L^{\beta-1}(y_1)) L(y_2) \prod_{j=1}^l (h\nu(dy_j)) \tag{4.18}
$$

with some constants κ_β^l. In case $k > 2$, this does not exceed $\tilde{a} f_{L^\beta}(h\nu) + \tilde{b}$ with some $\tilde{a}, \tilde{b}$, as by (ND$'$) the boundedness of $f_L(h\nu)$ implies the boundedness of $f_{\mathbf{1}}(h\nu)$. From this estimate, (4.17), and Gronwall's lemma, one obtains (4.16). In case $k = 2$ one sees that $\|Z^{h\nu}(t)\|$ is not involved in (4.18) and we get the same conclusion as for $k > 2$ using only the boundedness of $f_L(Z^{h\nu}(t))$ (and induction in β if necessarily). $\square$

We can now prove our last probabilistic result.

Theorem 4.3. *Under assumptions of Proposition* 4.1 *suppose additionally that the family of initial measures* $h\nu$ *converges* (*weakly*) *to a measure* $\mu \in \mathcal{M}^+(X)$

and that the initial conditions $h\nu$ are such that $\int (L^\beta)(x)h\nu(dx) \le C$ for all $h\nu$ and some $C > 0$ and an integer $\beta > 1$, the transition kernel P is $(1+L)^+$-bounded, is L-non-increasing and (ND$'$) *holds.*

Then there is a subsequence of the family of processes $Z^{h\nu}(t)$ defined by (3.4) *that weakly converges to a global non-negative solution μ_t of* (4.9) *such that*

$$\sup_{t\in[0,T]} \int (\mathbf{1} + L^\beta)(x)\mu_t(dx) \le C(T) \tag{4.19}$$

with some constant $C(T)$ for an arbitrary T.

Moreover, if P is not only L-non-increasing but also L-preserving, then the obtained solution μ_t is also L-preserving, i.e. for all t

$$\int L(x)\mu_t(dx) = f_L(\mu_t) = f_L(\mu_0) = \int L(x)\mu_0(dx).$$

Proof. Following the same strategy as in the proof of Theorem 4.1 leads to (4.10) that is implied by (4.11) and which is obvious for $(1+L)^+$- bounded kernels. It remains to show (4.13). Again one decomposes the domain of integration into the domains $\{\mathbf{z} = (z_1, \dots, z_m) : \max L(z_j) \ge K\}$ and its complement. The second part is dealt with precisely as in Theorem 4.1. Hence we need only to show that for any l, by choosing K arbitrary large, we can make

$$\int_{L(z_1)\ge K} (1 + L(z_1))\Big[\prod_{j=1}^{l} \mu_s(dz_j) + \prod_{j=1}^{l} Z^{h\nu}(s)(dz_j)\Big]$$

arbitrary small. But this integral does not exceed

$$K^{-(\beta-1)} \int (1 + L(z_1))(L(z_1))^{\beta-1}\Big[\prod_{j=1}^{l} \mu_s(dz_j) + \prod_{j=1}^{l} Z^{h\nu}(s)(dz_j)\Big],$$

which is finite and tends to zero as $K \to \infty$. Here we use (4.19) that follows directly from Proposition 4.2.

It remains to show that μ_t is L-preserving whenever the transition kernel $P(\mathbf{x}; d\mathbf{y})$ is L-preserving. As the approximations $Z^{h\nu}$ are then L-preserving, we only need to show that

$$\lim_{h\to 0} \int L(x)(\mu_t - Z^{h\nu}(t))(dx) = 0. \tag{4.20}$$

But this is done as above. Decomposing the domain of integration into two parts: $\{x : L(x) \le K\}$ and its complement, we observe that the required limit for the first integral is zero due to the weak convergence of $Z^{h\nu}(t)$ to μ_t, and on the second part both integrals from (4.19) can be made arbitrary small by choosing K large enough, due to the Proposition 4.2. $\square$

We shall now use the deterministic approach for proving existence in order to get rid of the nasty condition (ND$'$) and also to extend the previous result to (3.22). The necessity of additional conditions in the probabilistic approach from above arises from the need to show that the expectations of various almost surely finite objets is also finite.

Theorem 4.4. *Under the assumptions of Proposition* 4.1 *suppose additionally that* $\int(1+L^\beta)(x)\mu(dx)<\infty$ *for some* $\beta>1$*, the transition kernel* P *is* $(1+L)^+$*-bounded and* L*-non-increasing. Then there exist global non-negative solutions of* (4.9) *and of the corresponding integral version of* (3.22) *with the initial condition* μ *such that* (4.19) *holds for all* T.

Proof. We can assume that $\mu\neq 0$ (otherwise $\mu_t=0$ would be a required solution). Due to a priori low bounds for the solutions to (3.22) obtained in the above proof of Theorem 4.2, this case does not differ from (3.10), and we shall consider only the latter. We first approximate the transition kernel P by cutoff kernels P_n defined by the equation

$$\int g(\mathbf{y})P_n(\mathbf{z};d\mathbf{y})=\int \chi_{\{L^+(\mathbf{z})>n\}}(z)g(\mathbf{y})\chi_{\{L^+(\mathbf{y})>n\}}(y)P(\mathbf{z};d\mathbf{y}),$$

for arbitrary g, where χ_M denotes as usual the indicator function of the set M. It is easy to see that P_n enjoys the same property as P (it is still L-non-increasing and satisfies condition (ND)), but at the same time it is bounded, and hence the solution μ_t^n to the corresponding kinetic equations with initial condition μ exist by Theorem 4.1 (by Theorem 4.2 in case (3.22)). It is not difficult, of course, to construct direct deterministic proof of this fact. Moreover, as the evolution defined by P_n clearly does not change measures outside the compact region $\{y: L(y)\le n\}$, it follows that if $\int(1+L^\beta)(x)\mu(dx)<\infty$, then the same holds for μ_t for all t. Estimating $\Lambda f_{1+L^\beta}(\mu_t^n)$ by the same arguments as in the proof of Proposition 4.2 yields

$$\Lambda f_{1+L^\beta}(\mu_t^n)\le c\sum_{l=1}^k (L^\beta(y_1)+1)(L(y_2)+1)\prod_{j=1}^l \mu_t^n(dy_j)$$

with some constant c (independent of n), and consequently by Gronwall's lemma we get that for an arbitrary T

$$f_{1+L^\beta}(\mu_t^n)<C(T) \tag{4.21}$$

with some $C(T)$ for all $t\in[0,T]$ and all n. This implies that the family μ_t^n is compact for any t and any limiting point enjoys the same estimate (4.21). As the real-valued function $\int g(x)\mu_t^n(dx)$ is absolutely continuous (follows directly from (4.9)) for any $g\in C_c(X)$ choosing a countable dense sequence of such functions leads to the possibility to find a converging (in the space of continuous functions from $[0,T]$ to $\mathcal{M}^+(X)$, the latter taken with its weak topology) subsequence denoted also μ_t^n. It remains to show that its limit μ_t satisfies (4.9) by passing to the limit in the corresponding equations for μ_t^n. But this is obvious: all integrals outside the domain $\{y : L(y) < K\}$ can be made arbitrary small by choosing large K (because of (4.21)), and inside this domain the result follows from weak convergence. □

For conclusion, let us discuss the regularity of solutions constructed in the previous theorems showing in particular that they are in fact strong solutions of kinetic equations.

Theorem 4.5. *Let μ_t be a solution to* (4.9) (*or to the corresponding integral version of* (3.22)) *satisfying* (4.19) *with some $\beta > 1$. Let the transition kernel P be $(\mathbf{1} + L^\alpha)^+$-bounded for some $\alpha \in [0,1]$ and L-non-increasing. Then*

(i) (4.9) (*or the corresponding integral version of* (3.22)) *holds not only for $g \in C_c(X)$ but for any measurable g on X such that $|g| \le C(1 + L^{\beta-\alpha})$ with some constant C;*

(ii) *the function $\int g(x)\mu_t(dx)$ is an absolutely continuous function of t for all g from* (i) *and it is continuous for any measurable g such that $|g(x)| \le C(\mathbf{1} + L^\beta)(x)$;*

(iii) *for any g from* (i), *the function $\int g(x)\mu_t(dx)$ is a continuously differentiable function of t and* (3.10) *holds* (*respectively* (3.22));

(iv) *if the kernel P satisfies the additional condition*

$$\|P(\mu; z; \cdot) - P(\nu; z; \cdot)\| \le C(1 + L^\alpha)^+(z)\|(1 + L)(\mu - \nu)\| \tag{4.22}$$

(*for all finite measures μ, ν and some constant C*), *then the function $t \mapsto \mu_t$ is a continuously differentiable function in the sense of the total variation norm topology of $\mathcal{M}(X)$ and* (3.7) *holds in the strong sense, i.e. the derivative is again understood in the sense of the total variation norm of $\mathcal{M}(X)$.*

Proof. (i) One shows that (4.9) holds for all bounded measurable functions with a compact support by a straightforward limit argument. Hence, it is enough to prove the claim for measurable g such that $|g| \le L^{\beta-\alpha}$. This is obtained again by approximating g by functions with a compact support using the dominated convergence theorem and the estimate

$$\begin{aligned}\int g^+(\mathbf{y})P(\mu_s; \mathbf{z}; d\mathbf{y})\mu_s^{\tilde{\otimes}}(d\mathbf{z}) &\le \int (L^{\beta-\alpha})^+(\mathbf{y})P(\mu_s; \mathbf{z}; d\mathbf{y})\mu_s^{\tilde{\otimes}}(d\mathbf{z})\\ &\le \int (L^{\beta-\alpha})^+(\mathbf{z})(\mathbf{1} + L^\alpha)^+(\mathbf{z})\mu_s^{\tilde{\otimes}}(d\mathbf{z})\end{aligned}$$

when passing to the limit on both sides of (4.9).

(ii) It follows directly from (4.9) that the function $\int g(x)\mu_t(dx)$ is absolutely continuous for any g from (i). The required continuity for more general g is obtained again by approximating these g by functions with a compact support using the observation that the integrals of g and its approximations over the set $\{x : L(x) \geq K\}$ can be made uniformly arbitrary small for all μ_t, $t \in [0, T]$, by choosing K large enough.

(iii) For g from (i), equation (3.10) holds almost everywhere, which follows from (4.9). But as it follows from (ii), the integral

$$\int \phi(z_1, \dots, z_l) \prod_{j=1}^{l} \mu_t(dz_j)$$

is a continuous function of t for any symmetric measurable function ϕ on X^l such that

$$|\phi(z_1, \dots, z_l)| \leq C \prod_{j=1}^{l} (\mathbf{1} + L^\beta)(z_j)$$

(for all $|z_j| \leq K$ with an arbitrary K we can approximate these ϕ by products, and outside this domain all integrals are uniformly small). Consequently the r.h.s. of (3.10) is continuous, which implies the claim in (iii).

(iv) Choosing g to be characteristic functions of Borel sets, we get (3.7) set-wise. As

$$\|\mu_t - \mu_s\| = \sup_{g:|g|\leq 1} \int g(x)(\mu_t - \mu_s)(dx) = O(t - s)$$

uniformly for $0 \leq s \leq t \leq T$ with an arbitrary T, we conclude that $\|\mu_t\|$ is an absolutely continuous function of t in the sense of the norm topology on $\mathcal{M}(X)$. By the above reasoning we conclude that the norm continuity of the map $t \mapsto \mu_t$ implies the norm continuity of the map $t \mapsto L^\alpha(x)\mu_t(dx)$, which together with (4.22) implies the norm continuity of the r.h.s. of (3.7). To complete the proof of the theorem, it remains to notice that if $\dot{\mu}_t = \nu_t$ set-wise and if $t \mapsto \nu_t$ is continuous in the sense of the norm in $\mathcal{M}(X)$, then $\dot{\mu}_t = \nu_t$ holds in the sense of the norm topology. In fact,

$$\mu_t(B) = \mu_s(B) + (t - s)\nu_s(B) + \int_s^t (\nu_\tau(B) - \nu_s(B))\, ds,$$

and the norm of the integral on the r.h.s. is of order $o(t - s)$ as $t \to s$, whenever $\|\nu_t - \nu_s\| = o(1)$ as $t \to s$. □

To conclude this section let us notice that the regularity of the solution μ_t (the number of continuous derivatives) increases with the growth of β in (4.13).

As by Proposition 4.2 estimate (4.19) for μ_t follows from the corresponding estimate for μ_0, the regularity of μ_t depends on the rate of decay of μ_0 at infinity. For example, the following result holds.

Proposition 4.3. *If $f_{1+L^\beta}(\mu_0)$ is finite for all positive β and P does not depend explicitly on μ (e.g. no additional mean-field interaction is assumed), then the solution μ_t of* (3.10) *or* (3.22) *obtained by Theorems* 4.3–4.5 (*which is the unique solution by Theorem* 6.1 *below*) *is infinitely differentiable in t* (*with respect to the norm topology in $\mathcal{M}^+(X)$*).

5. Approximation by Markov chains

This short section is devoted to a possible approximation of the dynamics given by kinetic equations by means of Markov chains obtained by a discretization of X. For simplicity, we shall assume here that P does not depend on μ (i.e. there is no additional mean field interaction in the model). For a positive ω, a *marked ω-partition* of X, is a pair (X^ω, U^ω) where $X^\omega = \{M_j\}$, $j = 1, 2, \ldots$, is a countable subset of X and $U^\omega = \{U_j^\omega\}$ is a sequence of pairwise disjoint Borel subsets of X with the diameters not exceeding ω forming its partition, i.e. $X = \cup_{j=1}^\infty U_j$, such that $M_j \in U_j$ for all j. We shall write sometimes $U(M_j)$ for U_j that contains M_j. Clearly the countable set $X^\omega = \{M_1, M_2, \ldots\} \subset X$ is a ω-net of X, i.e. for any $x \in X$ there exists M_j such that $d(x, M_j) \le \omega$.

Remark 5.1. If $X = \mathbf{R}^d$, the lattice $\omega\mathbf{Z}^d$ is the most natural candidate for X^ω.

We shall say that a marked ω-partition of X is *L-admissible* if for all j

$$L(M_j) = \min\{L(y) : y \in U_j\}.$$

It is easy to see that L-admissible marked ω-partitions exist for any ω.

To an integral operator of type (3.3) on $B(S\mathcal{X})$ there corresponds a finite-difference operator $G_{\le k}^\omega$ on $B(S\mathcal{X}^\omega)$ (also of type (3.3)) defined by

$$G_{\le k}^\omega f(x_1, \ldots, x_n) = \sum_{I \subset \{1,\ldots,n\}} \sum_{m=0}^{\infty} \sum_{q_1,\ldots,q_m} P_{\mathbf{x}_I}^{q_1,\ldots,q_m}(\omega) \times (f(\mathbf{x}_{\bar{I}}, q_1, \ldots, q_m) - f(x_1, \ldots, x_m)), \tag{5.1}$$

$x_j, q_j \in X^\omega$, where

$$P_{x_1,\ldots,x_l}^{q_1,\ldots,q_m}(\omega) = P_m^l(x_1, \ldots, x_l; U(q_j) \times \ldots \times U(q_m)) \tag{5.2}$$

The corresponding scaled operator of type (3.4) has the form

$$\Lambda^{h,\omega} f(h\nu) = \frac{1}{h} \sum_{l=0}^{k} h^l \sum_{I \subset \{1,\ldots,n\}, |I|=l} \sum_{m=0}^{\infty} \sum_{q_1,\ldots,q_m} \Big[f\Big(h\nu - \sum_{i \in I} h\delta_{x_i}$$

$$+ h\delta_{y_1} + \ldots + h\delta_{y_m}\Big) - f(h\nu)\Big] P^{q_1,\ldots,q_m}_{\mathbf{x}_I}(\omega). \qquad (5.3)$$

Let us denote by $Z^{h\nu}_\omega$ the Markov chain on $\mathcal{M}^+_{h\delta}(X^\omega) \subset \mathcal{M}^+(X)$ specified by generator (5.3) and the initial condition $h\nu$. The L-admissibility of a marked ω-partition implies that if the assumptions of Proposition 4.1 hold for (3.4) and the corresponding transition kernel is L-non-increasing, then the discrete kernel (5.2) is also L-non-increasing and the assumptions of Proposition 4.1 hold also for (5.3). This implies, in particular, that $Z^{h\nu}_\omega$ are well defined, and moreover, as all estimates are uniform for $\omega \le 1$, the family of processes $Z^{h\nu}_\omega$, $h \in (0,1]$, $\omega \in (0,1]$ is tight.

The following result is an analogue of Theorem 4.3 for Markov chain approximations.

Proposition 5.1. *Under the assumptions of Theorem* 4.3 *there is a subsequence of the family of processes* $Z^{h\nu}_\omega$ *that weakly converges* (*as* $h \to 0$, $\omega \to 0$) *to a global weak solution* μ_t *of the corresponding kinetic equation, i.e.* μ_t *satisfies* (3.10).

Proof. This is the same as the proof of Theorem 4.3 (supplemented by Proposition 4.2). The only difference is the necessity to prove

$$|\Lambda^{h,\omega} f_g(Z^{h\nu}_\omega(s)) - \Lambda f_g(Z^{h\nu}_\omega(s))| \to 0, \quad h \to 0, \quad \omega \to 0, \qquad (5.4)$$

instead of (4.10). But as the terms of type (4.11) are small for small h, to get (5.4) one only needs to show that the difference between

$$\sum_{l=1}^k \frac{1}{l!} \int [g^+(\mathbf{y}) - g(z_1) - \ldots - g(z_l)] P(z_1,\ldots,z_l; d\mathbf{y}) \prod_{j=1}^l (h\eta(dz_j))$$

and

$$\sum_{l=1}^k \frac{1}{l!} \sum_{m=0}^\infty \sum_{q_1,\ldots,q_m} [g(q_1) + \ldots g(q_m) - g(z_1) - \ldots - g(z_l)] P^{q_1,\ldots,q_m}_{z_1,\ldots,z_l} \prod_{j=1}^l (h\eta(dz_j))$$

tends to zero as $h \to 0$. But this holds by (5.2), as g has a compact support.

6. Uniqueness and continuous dependence on initial data

Theorem 6.1. *Suppose the assumptions of Theorem* 4.4 *hold,* P *is* $(1 + L^\alpha)^+$*-bounded for some* $\alpha \in [0,1]$ *such that* $\beta \ge \alpha + 1$, *and* (4.22) *holds. Then there exists a unique non-negative solution* μ_t *to* (3.10) (*and similarly to* (3.22)) *satisfying* (4.19) *and a given initial condition* μ_0 *such that*

$$\int (1 + L^\beta)(x)\mu_0(dx) < \infty.$$

This μ_t is a strong solution of the corresponding kinetic equation, i.e. it satisfies (3.7) (*or the corresponding strong analog of* (3.22)), *where the derivative is understood in the sense of the norm in $\mathcal{M}(X)$.*

Moreover, for any two solutions μ_t and ν_t of (3.10) (or (3.22)) *satisfying* (4.19) *with initial conditions μ_0 and ν_0 one has*

$$\int (\mathbf{1}+L)|\mu_t-\nu_t|\,(dx) \le ae^{at}\int (\mathbf{1}+L)|\mu_0-\nu_0|\,(dx) \tag{6.1}$$

for some constant a uniformly for all $t\in[0,T]$.

Remark 6.1. Of course, it follows from this theorem that under the assumptions of Theorem 4.3 the whole family of processes $Z^{h\nu}(t)$ from Theorem 4.3 (not just its subsequence) with the weak limit μ_0 of $Z^{h\nu}(0)$ converges weakly to μ_t; also the family of Markov chains $Z^{h\nu}_\omega$ (defined in Section 5) converges weakly, as $h\to 0$, $\omega\to 0$, to the same limit.

Proof. By the previous results we only need to prove (6.1). Again we shall consider only equation (3.10), as (3.22) can be effectively included in this case due to the a priori lower bounds on the solutions obtained in the proof of Theorem 4.2. As by Theorem 4.2 (iv), μ_t and ν_t are solutions of (3.7) with continuous right hand sides, Lemma A from the Appendix (see also remark after this lemma) can be applied to the measure $(1+L)(x)(\mu_t-\nu_t)(dx)$. Consequently, denoting by f_t a version of the density of $\mu_t-\nu_t$ with respect to $|\mu_t-\nu_t|$ from Lemma A yields

$$\begin{aligned}\int (1+L)(x)|\mu_t-\nu_t|(dx) &= \|(1+L)(\mu_t-\nu_t)\|\\ &= \int (1+L)(x)|\mu_0-\nu_0|(dx)\\ &\quad+\int_0^t ds\int_X f_s(x)(1+L)(x)(\dot\mu_s-\dot\nu_s)(dx).\end{aligned} \tag{6.2}$$

By (3.7) the last integral in (6.2) equals

$$\begin{aligned}&\int_0^t ds\sum_{l=1}^k\int\int \big([f_s(1+L)]^+(\mathbf{y})-[f_s(1+L)]^+(\mathbf{z})\big)\\ &\quad\times\Big(P(\mu;\mathbf{z};d\mathbf{y})\prod_{j=1}^l\mu_s(dz_j)-P(\nu;\mathbf{z};d\mathbf{y})\prod_{j=1}^l\nu_s(dz_j)\Big)\\ &\quad=\int_0^t ds\sum_{l=1}^k\int\int \big([f_s(1+L)]^+(\mathbf{y})-[f_s(1+L)]^+(\mathbf{z})\big)\end{aligned}$$

$$\times \Big[P(\mu;\mathbf{z};d\mathbf{y}) \sum_{j=1}^{l} \prod_{i=1}^{j-1} \nu_s(dz_i)(\mu_s-\nu_s)(dz_j) \prod_{i=j+1}^{l} \mu_s(dz_i)$$

$$+ (P(\mu;\mathbf{z};d\mathbf{y}) - P(\nu;\mathbf{z};d\mathbf{y})) \prod_{j=1}^{l} \nu_s(dz_j)\Big]. \tag{6.3}$$

Let us pick up arbitrary $l \leq k$ and $j \leq l$ and estimate the corresponding term of the sum in the square bracket of (6.3). We have

$$\int\int \big([f_s(1+L)]^+(\mathbf{y}) - [f_s(1+L)]^+(\mathbf{z})\big) P(\mu;\mathbf{z};d\mathbf{y})$$

$$\times (\mu_s-\nu_s)(dz_j) \prod_{i=1}^{j-1} \nu_s(dz_i) \prod_{i=j+1}^{l} \mu_s(dz_i)$$

$$= \int\int \big([f_s(1+L)]^+(\mathbf{y}) - [f_s(1+L)]^+(\mathbf{z})\big) P(\mu;\mathbf{z};d\mathbf{y})$$

$$\times f_s(z_j)\,|\mu_s-\nu_s|\,(dz_j) \prod_{i=1}^{j-1} \nu_s(dz_i) \prod_{i=j+1}^{l} \mu_s(dz_i). \tag{6.4}$$

As L is non-increasing by $P(\mathbf{z};d\mathbf{y})$,

$$\big([f_s(1+L)]^+(\mathbf{y}) - [f_s(1+L)]^+(\mathbf{z})\big) f_s(z_j)$$

$$\leq (1+L)^+(\mathbf{y}) - f_s(z_j)[f_s(1+L)]^+(\mathbf{z})$$

$$\leq m_0 + k + L^+(\mathbf{z}) - L(z_j) - \sum_{i\neq j} f_s(z_j) f_s(z_i) L(z_i)$$

$$\leq m_0 + k + 2\sum_{i\neq j} L(z_i),$$

where the constant m_0 is from condition (ND). Hence (6.4) does not exceed

$$\int (m_0 + k + 2\sum_{i\neq j} L(z_i))(k + L^\alpha(z_j)$$

$$+ \sum_{i\neq j} L^\alpha(z_i))\,|\mu_s-\nu_s|\,(dz_j) \prod_{i=1}^{j-1} \nu_s(dz_i) \prod_{i=j+1}^{l} \mu_s(dz_i).$$

Consequently, as $1+\alpha \leq \beta$ and (4.19) holds, and since the second term in the square bracket of (6.3) can be estimated by (4.22), it follows that the integral (6.3) does not exceed

$$a(r,T) \int_0^t ds \int (1+L)(x)\,|\mu_t-\nu_t|\,(dx)$$

with some constant $a(r,T)$, which implies (6.1) by Gronwall's lemma. Proof of the theorem is complete. □

As a consequence, we shall obtain now a version of the propagation of chaos property for interacting particle systems considered in Section 3. In general, this property means (see e.g. [35]) that the moment measures of some random measures tend to the product measures when passing to a certain limit. The moment measures μ_t^m of the jump processes $Z^{h\nu}(t)$ from Theorem 4.1 are defined (see e.g. [11]) as

$$\mu_{t,h}^m(dx_1 \dots dx_m) = \mathsf{E}\left(Z^{h\nu}(t)(dx_1)\dots Z^{h\nu}(t)(dx_m)\right). \tag{6.5}$$

Theorem 6.2. *Under the conditions of Theorem* 4.3 *suppose* P *is* $(1+L^\alpha)^+$-*bounded for some* $\alpha \in [0,1]$ *such that* $\beta \geq \alpha+1$, *and* (4.22) *holds. Let the family of initial measures* $h\nu = h\nu(h)$ *converge weakly to a certain measure* μ_0 *as* $h \to 0$. *Then for any* $m = 1,2,\dots$, *the moment measures* $\mu_{t,h}^m$ *converge weakly to the product measure* $\mu_t^{\otimes m}$.

Proof. By Theorem 6.1, for any $g \in C_c(SX^m)$ the random variables

$$\eta_h = \int g(x_1,\dots,x_m) Z^{h\nu}(t)(dx_1)\dots Z^{h\nu}(t)(dx_m)$$

converge almost surely to the $\int g(x_1,\dots,x_m)\prod_{j=1}^m \mu_t(dx_j)$. To conclude that the expectations of η_h converge to its point-wise deterministic limit, one only need to show that the variances of these random variable are uniformly bounded. Consequently, to complete the proof one needs to show that $\mathsf{E}(\|Z^{h\nu}(t)\|^r)$ are bounded for any positive integer r uniformly for $t\in[0,T]$ with arbitrary $T>0$. But this is done like in the proof of Proposition 4.1 (i). Namely,

$$\Lambda^h(f_{\mathbf{1}})^r(h\nu) \leq h^r \sum_{i=1}^n \sum_{m=0}^\infty \int ((n+m-1)^r - n^r) P(h\nu; x_i; dy_1\dots dy_m),$$

for $\nu = \delta_{x_1} + \dots + \delta_{x_n}$ and hence

$$\Lambda^h f_{\mathbf{1}}^r(h\nu) \leq m_0 r h^r (n+m_0)^{r-1} \sum_{i=1}^n (\mathbf{1}+L)(x_i) \leq K(f_{\mathbf{1}}^r(h\nu)+1),$$

as the number of particles created in one go is bounded by m_0 and $f_L(h\nu)$ are uniformly bounded. Hence it follows by Gronwall's lemma that $\mathsf{E}\,\|Z^{h\nu}(t)\|^r$ are uniformly bounded, which completes the proof. □

A. Appendix

Lemma A.1. *Let Y be a measurable space and the mapping $t \mapsto \mu_t$ from $[0,T]$ to $\mathcal{M}(Y)$ is continuously differentiable in the sense of the norm in $\mathcal{M}(Y)$ with a (continuous) derivative $\dot{\mu}_t = \nu_t$. Let σ_t denote a density of μ_t with respect to its total variation $|\mu_t|$, i.e. the class of measurable functions taking three values $-1, 0, 1$ and such that $\mu_t = \sigma_t|\mu_t|$ and $|\mu_t| = \sigma_t\mu_t$ almost surely with respect to $|\mu_t|$. Then there exists a measurable function $f_t(x)$ on $[0,T] \times Y$ such that f_t is a representative of class σ_t for any $t \in [0,T]$ and*

$$\|\mu_t\| = \|\mu_0\| + \int_0^t ds \int_Y f_s(y)\nu_s(dy). \tag{A.1}$$

Proof. Step 1. As μ_t is continuously differentiable, $\|\mu_t - \mu_s\| = O(t-s)$ uniformly for $0 \le s \le t \le T$. Hence $\|\mu_t\|$ is an absolutely continuous real-valued function. Consequently, this function has almost everywhere on $[0,T]$ a derivative, say ω_s, and has an integral representation $\|\mu_t\| = \|\mu_0\| + \int_0^t \omega_s \, ds$ valid for all $t \in [0,T]$. It remains to calculate ω_s. To simplify this calculation, we observe that as the right and left derivatives of an absolutely continuous function coincide almost everywhere (Lebesgue theorem), it is enough to calculate only the right derivative of $\|\mu_t\|$. Hence from now on we shall consider only the limits $t \to s$ with $t \ge s$.

Step 2. For an arbitrary measurable $A \subset Y$ and an arbitrary representative σ_t of the density, we have

$$\begin{aligned} O(t-s) &= \int_A |\mu_t|(dy) - \int_A |\mu_s|(dy) = \int_A \sigma_t\mu_t(dy) - \int_A \sigma_s\mu_s(dy) \\ &= \int_A (\sigma_t - \sigma_s)\mu_s(dy) + \int_A \sigma_t(\mu_t - \mu_s)(dy). \end{aligned}$$

As the second term here is also of order $O(t-s)$, we conclude that the first term is of order $O(t-s)$ uniformly for all A and s, t. Hence $\sigma_t \to \sigma_s$ almost surely with respect to $|\mu_s|$ as $t \to s$. As σ_t takes only three values $(0, 1, -1)$, it follows that $\dot{\sigma}_s(x)$ exists and vanishes for almost all x with respect to $|\mu_s|$.

Remark A.1. Writing now formally

$$\frac{d}{dt}\|\mu_t\| = \frac{d}{dt}\int_Y \sigma_t\mu_t(dy) = \int_Y \dot{\sigma}_t\mu_t(dy) + \int_Y \sigma_t\dot{\mu}_t(dy)$$

and noticing that the first term here vanishes (by Step 2) yields (A.1). However, this formal calculation can not be justified for an arbitrary choice of σ_t.

Step 3. Let us choose now an appropriate representative of σ_t. For this purpose let us perform the Lebesgue decomposition of ν_t into the sum $\nu_t = \nu_t^s + \nu_t^a$ of a singular and an absolutely continuous measure with respect to $|\mu_t|$. Let η_t be the density of ν_t^s with respect to its total variation measure $|\nu_t^s|$, i.e. the class of functions taking values $0, 1, -1$ and such that $\nu_t^s = \eta_t |\nu_t^s|$ almost surely with respect to $|\nu_t^s|$. Now let us pick up an f_t from the intersection $\sigma_t \cap \eta_t$, i.e. f is a representative of both density classes simultaneously. Such a choice is possible, because μ_t and ν_t^s are mutually singular. From this definition of f_t and from Step 2 it follows that $f_t \to f_s$ as $t \to s$ (and $t \geq s$) almost surely with respect to both μ_s and ν_s. In fact, let B be either a positive part in the Hahn decomposition of ν_s^s or any its measurable subset. Then $f_s = 1$ on B. Moreover,

$$\mu_t(B) = (t-s)\nu_s^s(B) + o(t-s), \quad t - s \to 0,$$

and is positive and hence $f_t = 1$ on B for t close enough to s.

Step 4. By definition,

$$\begin{aligned} \frac{d}{ds}\|\mu_s\| &= \lim_{t\to s} \frac{\|\mu_t\| - \|\mu_s\|}{t-s} \\ &= \lim_{t\to s} \int \frac{f_t - f_s}{t-s}\mu_s(dy) + \lim_{t\to s} \int f_t \frac{\mu_t - \mu_s}{t-s}(dy) \end{aligned} \tag{A.2}$$

(if both limits exist, of course). It is easy to see that the second limit here always exists and equals $\int f_s \nu_s(dy)$. In fact,

$$\begin{aligned} \int f_t \frac{\mu_t - \mu_s}{t-s}(dy) &= \int f_s \frac{\mu_t - \mu_s}{t-s}(dy) \\ &\quad + \int (f_t - f_s)\Big(\frac{\mu_t - \mu_s}{t-s} - \nu_s\Big)(dy) + \int (f_t - f_s)\nu_s(dy), \end{aligned}$$

and the limit of the first integral equals $\int f_s \nu_s(dy)$, the second integral is of order $o(t-s)$ by the definition of the derivative and hence vanishes in the limit, and the limit of the third integral is zero because (due to our choice of f_t in Step 3) $f_t - f_s \to 0$ as $t \to s$ $(t \geq s)$ almost surely with respect to ν_s. Consequently, to complete the proof it remains to show that the first term on the r.h.s. of (A.2) vanishes.

Remark A.2. As we showed in Step 2, the function $(f_t - f_s)/(t-s)$ under the integral in this term tends to zero almost surely with respect to μ_s, but unfortunately this is not enough to conclude that the limit of the integral vanishes, and consequently an additional argument is required.

Step 5. In order to show that the first term in (A.2) vanishes, it is enough to show that the corresponding limits of the integrals over the sets A^+ and A^-

vanish, where $Y = A^+ \cup A^- \cup A^0$ is the Hahn decomposition of Y with respect to the measure μ_s. Let us consider only A^+ (A^- is considered similarly). Hence, as $f_s = 1$ on A^+ almost surely, we need to show that

$$\lim_{t \to s} \int_A \frac{f_t - 1}{t - s} \mu_s(dy) = 0 \tag{A.3}$$

where A is a measurable subset of Y such that $(\mu_s)|_A$ is a positive measure. Using now the Lebesgue decomposition of $(\nu_s)|_A$ into the sum of a singular and absolutely continuous parts with respect to μ_s, we can and will reduce the discussion to the case when ν_s is absolutely continuous with respect to μ_s on A.

Introducing the set $A_t = \{y \in A : f_t(y) \leq 0\}$ one can clearly replace A by A_t in (A.3). Consequently, to get (A.3) it is enough to show that $\mu_s(A_t) = o(t-s)$ as $t \to s$. This will be done in the next final step.

Step 6. From the definition of A_t it follows that $\mu_t(B) \leq 0$ for any $B \subset A_t$ and hence

$$\mu_s(B) + (t-s)\nu_s(B) + o(t-s) \leq 0, \tag{A.4}$$

where $o(t-s)$ is uniform, i.e. $\|o(t-s)\|/(t-s) \to 0$ as $t \to s$. Notice first that if $A_t = B_t^+ \cup B_t^- \cup B_t^0$ is the Hahn decomposition of A_t on the positive, negative and zero parts of the measure ν_s, then $\mu_s(B_t^+ \cup B_t^0) = o(t-s)$ uniformly (as it follows directly from (A.4)), and consequently we can and will reduce our discussion to the case when ν_s is a negative measure on A. In this case (A.4) implies that

$$\mu_s(A_t) \leq (t-s)(-\nu_s)(A_t) + o(t-s)$$

and it remains to show that $\nu_s(A_t) = o(1)_{t \to s}$. To see this we observe that for any s, Y has the representation $Y = \cup_{n=0}^\infty Y_n$, where $|\mu_s|(Y_0) = 0$ and

$$Y_n = \{y \in Y : f_t = f_s \quad \text{for } |t-s| \leq 1/n\}.$$

Clearly $Y_n \subset Y_{n+1}$ for any $n \neq 0$ and $A_t \subset Y \setminus Y_n$ whenever $t - s \leq 1/n$. Hence A_t are subsets of a decreasing family of sets with an intersection of μ_s-measure zero. As ν_s is absolutely continuous with respect to μ_s the same holds for ν_s and hence $\nu_s(A_t) = o(1)_{t \to s}$, which completes the proof of the lemma. □

Remark A.3. Suppose the assumptions of Lemma A hold and $L(y)$ is a measurable, non-negative and everywhere finite function on Y such that $\|L\mu_s\|$ and $\|L\nu_s\|$ are uniformly bounded for $s \in [0,t]$. Then (A.1) holds with $L\mu_t$ and $L\nu_t$ instead of μ_t and ν_t respectively. In fact, though $s \mapsto L\nu_s$ may be discontinuous in the sense of norm, one can write the required identity first with the space Y_m instead of Y, where $Y_m = \{y : L(y) \leq m\}$, and then pass to the limit as $m \to \infty$.

Lemma A.2. *Suppose X is a locally compact metric space, L is a non-negative continuous function on it and a, b are two positive constants. Then the solution to the measure-valued ordinary differential equation*

$$\dot{\mu}(dx) = -aL(x)\mu(dx) - b\,\frac{\int L(u)\mu(du)}{\int \mu(du)}\,\mu(dx) \tag{A.5}$$

on (positive) measures μ on a X with initial condition μ_0 equals

$$\mu_t(dx) = \exp\{-atL(x)\}\mu_0(dx)\exp\left\{-b\int_0^t \frac{\int L(u)\exp\{-asL(u)\}\mu_0(du)}{\int \exp\{-asL(u)\}\mu_0(du)}\,ds\right\} \tag{A.6}$$

and in particular it never vanishes whenever $\mu_0 \neq 0$.

Proof. This is obtained by simple calculus and we just sketch it. Clearly it is enough to get this solution for measures having density $y(x)$ with respect to some reference measure dx on X. For (non-negative) functions $y(x)$ equation (A.5) becomes

$$\dot{y}_t(x) = -aL(x)y_t(x) - b\,\frac{\int L(u)y_t(u)du}{\int y_t(u)du}\,y_t(x). \tag{A.7}$$

Dividing by y yields

$$\frac{d}{dt}\ln y_t(x) = -aL(x) + g(t)$$

with some function $g(t)$ depending only on t (and not on x). Consequently

$$y_t(x) = \exp\{-atL(x)\}\beta(x)\gamma(t)$$

with some functions β and γ. Taking into account the initial condition leads to the representation

$$y_t(x) = \exp\{-atL(x)\}y_0(x)\gamma(t)$$

with some $\gamma(t)$ such that $\gamma(0) = 1$. Substituting this function again into equation (A.7) yields the linear equation

$$\dot{\gamma}(t) = -b\gamma(t)\,\frac{\int L(u)\exp\{-atL(u)\}y_0(u)\,du}{\int \exp\{-atL(u)\}y_0(u)\,du}$$

for γ. Solving this equation yields (A.6). □

Acknowledgements

I am thankful to V.P. Belavkin who first brought my attention to the theory of interacting particle systems and to J. Norris for organizing two stimulating conferences on coagulation problems in Oberwolfach and Cambridge.

References

[1] O. Arino and R. Rudnicki (2004) Phytoplankton dynamics. *Comtes Rendus Biologies* **327**, 961–969.

[2] J.M. Ball and J. Carr (1990) The discrete coagultion-fragmentation equations: existence, uniqueness and density conservation. *J. Stat. Phys.* **61**, 203–234.

[3] V.P. Belavkin (1988) Quantum branching processes and nonlinear dynamics of multi-quantum systems. *Dokl. Acad. Nauk SSSR* **301** (6), 1348–1352. English tanslation: *Sov. Phys. Dokl.*, 1988, **33** (8), 581–582.

[4] V. Belavkin and V. Kolokoltsov (2001) On general kinetic equation for many particle systems with interaction, fragmentation and coagulation. Preprint 16/01 (2001). Department of Computing and Mathematics Nottingham Trent University. To appear in *Proc. Royal Society.*

[5] V.P. Belavkin and V.P. Maslov (1977) Uniformization method in the theory of nonlinear hamiltonian systems of Vlasov and Hartree type. *Teoret. i Matem. Fizika* **33** (1), 17–31. English translation: *Theor. Math. Phys.* **43** (3), 852–862.

[6] P. Brémaud (1981) *Point Processes and Queues.* Springer Series in Statistics, Springer.

[7] J. Carr (1992) Asymptotic behavior of solutions to the coagulation-fragmentation equations I. The strong fragmentation case. *Proc. Royal Soc. Edinburgh, Sect. A* **121**, 231–244.

[8] Mu Fa Chen (1992) *From Markov Chains to Non-Equilibrium Particle Systems.* World Scientific.

[9] V. Corradi and R. Sarin (2000) Continuous approximations of stochastic evolutionary game dynamic. *J. Econ. Theory* **94**, 163–191.

[10] Z. Cheng and S. Redner (1988) Scaling theory of fragmentation. *Phys. Rev. Lett.* **60**, 2450–2453.

[11] D. Dawson (1993) *Measure-Valued Markov Processes.* In: *Ecole d'Eté de Probabilités de Saint-Flour* XXI–1991, P.L. Hennequin (ed.), Lect. Notes Math. **1541**, 1–260.

[12] P.B. Dubovski (1999) A "triangle" of interconnected coagulation models. *J. Phys. A: Math. Gen.* **32**, 781–793.

[13] P.B. Dubovskii and I.W. Stewart (1996) Existence, uniqueness and mass conservation for the coagulation-fragmentation equation. *Math. Method Appl. Sci.* **19**, 571–591.

[14] E. Dynkin (1994) *An Introduction to Branching Measure-Valued Processes.* CRM monograph series **6**, AMS, Providence RI.

[15] S.N. Ethier and T.G. Kurtz (1986) *Markov Processes. Characterization and Convergence.* John Wiley & Sons.

[16] M.Z. Guo, G.Z. Papanicolau and S.R.S. Varadhan (1988) Nonlinear diffusion limit for a system with nearest neighbor interactions. *Commun. Math. Phys.* **118**, 31–59.

[17] J. Hofbauer and K. Sigmund (1998) *Evolutionary Games and Population Dynamics*. Cambridge University Press.

[18] I. Jeon (1998) Existence of Gelling Solutions for coagulation-fragmentation equations. *Commun. Math. Phys.* **194**, 541–567.

[19] V. Kolokoltsov (2003) Measure-valued limits of interacting particle systems with k-nary interactions I. *Probab. Theory Relat. Fields* **126**, 364–394.

[20] V. Kolokoltsov (2004) Measure-valued limits of interacting particle systems with k-nary interactions II. *Stochastics and Stochastics Reports* **76** (1), 45–58.

[21] V. Kolokoltsov (2004) Hydrodynamic limit of coagulation-fragmentation type models of k-nary interacting particles. *J. Stat. Phys.* **115** (5/6), 1621–1653.

[22] V. Kolokoltsov (2004) On Markov processes with decomposable pseudo-differential generators. *Stochastics and Stochastics Reports* **76:1**, 1–44.

[23] V. Kolokoltsov (2003) On extension of mollified Boltzmann and Smoluchovski equations to particle systems with a k-nary interaction. *Russian J. Math. Phys.* **10** (3), 268–295.

[24] V. Kolokoltsov (2000) Symmetric stable laws and stable-like jump-diffusions. *Proc. London Math. Soc.* **3** (80), 725–768.

[25] M. Kostoglou and A.J. Karabelas (2000) A study of the nonlinear breakage equations: analytical and asymptotic solutions. *J. Phys. A* **33**, 1221–1232.

[26] M. Lachowicz, Ph. Laurencot and D. Wrzosek (2003) On the Oort – Hulst – Savronov coagulation equation and its relation to the Smoluchowski equation. *SIAM J. Math. Anal.* **34**, 1399–1421.

[27] C. Landim, J. Quastel, M. Salmhofer and H.T. Yau (2004) Superdiffusivity of asymmetric exclusion process in dimensions one and two. *Commun. Math. Phys.* **244** (3), 455–481.

[28] P. Laurencot and S. Mischler (2002) From the discrete to the continuous coagulation-fragmentation equations. *Proc. Royal Soc. Edinburgh, Sect. A* **132** (5), 1219–1248.

[29] P. Laurencot and D. Wrzosek (2001) The discrete coagulation equations with collision breakage. *J. Stat. Phys.* **104** (1/2), 193–220.

[30] J. Norris (1999) Smoluchovski's coagulation equation: uniqueness, nonuniqueness and a hydrodynamic limit for the stochastic coalescent. *Ann. Prob.* **9** (1), 78–109.

[31] J. Norris (2000) Cluster coagulation. *Commun. Math. Phys.* **209**, 407–435.

[32] A.J. Povzner (1962) About the Boltzmann equation in kinetic gas theory. *Mat. Sbornik* **58**, 65–86.

[33] L. Rass and J. Radcliffe (2003) *Spatial Deterministic Epidemics*. Mathematical Surveys and Monographs **102**, AMS.

[34] V.S. Safronov (1969) *Evolution of the Pre-Planetary Cloud and the Formatio of the Earth and Planets*. Moscow, Nauka, (in Russian). English translation: Israel Program for Scientific Translations, Jerusalem, 1972.

[35] A. Sznitman (1991) Topics in propagation of chaos. In: *Ecole d'Eté de Probabilités de Saint-Flour* XIX–1989. Lect. Notes Math. **1464**, 167–255.

[36] J.W. Weibull (1995) *Evolutionary Game Theory*. The MIT Press.