NONLINEAR MARKOV SEMIGROUPS AND INTERACTING LÉVY TYPE PROCESSES

Vassili N. Kolokoltsov*

November 21, 2006

To appear in Journal of Statistical Physics.

Abstract

Semigroups of positivity preserving linear operators on measures of a measurable space X describe the evolutions of probability distributions of Markov processes on X. Their dual semigroups of positivity preserving linear operators on the space of measurable bounded functions $B(X)$ on X describe the evolutions of averages over the trajectories of these Markov processes. In this paper we introduce and study the general class of semigroups of non-linear positivity preserving transformations on measures that is non-linear Markov or Feller semigroups. An explicit structure of generators of such groups is given in case when X is the Euclidean space $\mathbf{R}^d$ (or more generally, a manifold) showing how these semigroups arise from the general kinetic equations of statistical mechanics and evolutionary biology that describe the dynamic law of large numbers for Markov models of interacting particles. Well posedness results for these equations are given together with applications to interacting particles: dynamic law of large numbers and central limit theorem, the latter being new already for the standard coagulation-fragmentation models.

Key words. Positivity preserving measure-valued evolutions, conditionally positive operators, Markov models of interacting particles, dynamic law of large numbers, normal fluctuations, rate of convergence, kinetic equations, interacting stable jump-diffusions, Lévy type processes, coagulation-fragmentation.

Running Head: Nonlinear Markov semigroups.

1 Introduction

1.1 Aims of the paper

An important class of Markov semigroups is given by the so called Feller semigroups, i.e. the semigroups of strongly continuous operators T_t , $t \geq 0$, on $C_\infty(X)$ (the space

*Department of Statistics, University of Warwick, Coventry CV4 7AL UK. Email: v.kolokoltsov@warwick.ac.uk

of continuous functions on a locally compact topological space X vanishing at infinity) such that $0 \leq u \leq 1$ implies $0 \leq T_t u \leq 1$ for all $t \geq 0$. In a wider sense, by Feller semigroups one understands the semigroups of positivity preserving linear operators in the space of bounded continuous functions $C(X)$, our basic reference for Feller semigroups being [28]. It is easy to see that if the evolution on $C_\infty(X)$ given by the equation $\dot{f} = Af$ with a linear (possibly unbounded) operator A in $C_\infty(X)$ preserves positivity, then A is conditionally positive, i.e. $(Ag)(x) \geq 0$ whenever a function g belongs to the domain of A, is non-negative and vanishes at x. If X is the Euclidean space $\mathbf{R}^d$ or a manifold, the Courrège theorem [12] states that a linear operator A in $C_\infty(X)$ whose domain contains the space $C_c^2(X)$ of two times continuously differentiable functions with a compact support is conditionally positive if and only if it has the Levy-Khintchine form with variable coefficients, i.e. if

$$\begin{aligned} Ag(x) =& a(x)g(x) + (b(x), \nabla)g(x) + \frac{1}{2}(c(x)\nabla, \nabla)g(x) \\ &+ \int (g(x+y) - g(x) - \chi(y)(y, \nabla)g(x))\nu(x; dy), \end{aligned} \tag{1}$$

where $c(x)$ is a positive definite $d \times d$-matrix, $b(x)$ is a d-vector, $\chi(y)$ is some bounded non-negative function with a compact support that equals one in a neighborhood of the origin, and $\nu(x, \cdot)$ is a Lévy measure, i.e. a positive Borel measure such that

$$\int \min(1, |y|^2)\nu(x; dy) < \infty, \qquad \nu(x)(\{0\}) = 0. \tag{2}$$

For what follows the main role will belong to the dual formulation of Feller semigroups. Namely, a Feller semigroup T_t on $C_\infty(X)$ clearly gives rise to a dual positivity preserving semigroup T_t^* on the space $\mathcal{M}(X)$ of bounded Borel measures on X through the duality identity $(T_t f, \mu) = (f, T_t^* \mu)$, where the pairing (f, μ) is given by the integration, of course. If A is the generator of T_t, then $\mu_t = T_t^* \mu$ can be characterized by the equation in the weak form

$$\frac{d}{dt}(g, \mu_t) = (Ag, \mu_t) = (g, A^* \mu_t), \tag{3}$$

where A^* is the adjoint to A, that holds for all g from the domain of A.

Here we shall deal with nonlinear analogs of (3):

$$\frac{d}{dt}(g, \mu_t) = \Omega(\mu_t) g, \tag{4}$$

where Ω is a nonlinear transformation from a dense domain of $\mathcal{M}(X)$ to the space of linear functionals on $C(X)$. Of special interest are the equations of the form

$$\frac{d}{dt}(g, \mu_t) = (A(\mu_t)g, \mu_t) = (g, A^*(\mu_t)\mu_t), \tag{5}$$

where $A(\mu)$ is a nonlinear mapping from a dense domain of $\mathcal{M}(X)$ to (possibly unbounded) linear operators in $C_\infty(X)$, and even more specifically the equations with A in (5) depending polynomially on μ, i.e. the equations of the form

$$\frac{d}{dt}(g, \mu_t) = \sum_{k=1}^{K} \int \cdots \int (A_k g)(x_1, \ldots, x_k) \mu_t(dx_1) \cdots \mu_t(dx_k), \tag{6}$$

where each A_k is a (possibly unbounded) operator from a dense subspace D of $C_\infty(X)$ to the space $C^{sym}(X^k)$ of symmetric continuous bounded functions of k variables from X .

Our aim is two folds: 1) pure analytic problem (Sections 2-4) - to introduce an appropriate analog of the notion of conditional positivity for the "generators" Ω and A in (4), (5) that is characteristic for positivity preserving measure-valued evolutions, to give an explicit general structure of conditionally positive generators (a nonlinear analog of the Courrège theorem) and to provide basic well posedness results for the corresponding evolution equations (5), (6); 2) application (Sections 5-7) - to develop analytic tools for proving the law of large numbers and the central limit theorems together with precise rates of convergence for a wide class of interacting Markov processes with pseudo-differential generators, i.e. for interacting Lévy type processes.

1.2 Motivation: Markov models of interacting particles

A system of mean field interacting particles (or Markov processes) with an underlying motion described by generator (1) is a system of N particles in X moving according to a generator of type (1), where the coefficients a_j , b_j , c_j , γ_j specifying the motion of every j -th particle depend not only on the position x_j of this particle, but also on the point measure $\mu = h(\delta_{x_1} + \cdots + \delta_{x_N})$ on X (mean field) specified by the position of all other particles $x_1, \ldots, x_N$ of the system (h being a positive scaling parameter, e.g. $h = 1/N$), i.e. $a_j(x_j) = a(x_j, \mu)$, $b_j(x_j) = b(x_j, \mu)$, $c_j(x_j) = c(x_j, \mu)$, $\nu_j(x_j) = \nu(x_j, \mu)$. It is well known (and can be easily seen) that passing to the mean field or McKean - Vlasov limit in this system (a scaling limit with the number of particles N tending to infinity in such a way that the scaled measures $h(\delta_{x_1} + \cdots + \delta_{x_N})$ tend to some finite measures) leads formally to a measure-valued dynamics on X described by equation (5) with

$$A(\mu)g(x) = a(x, \mu)g(x) + (b(x, \mu), \nabla)g(x) + \frac{1}{2}(c(x, \mu)\nabla, \nabla)g(x) + \int (g(x + y) - g(x) - \chi(y)(y, \nabla)g(x))\nu(x, \mu; dy) \quad (7)$$

that specifies the deterministic dynamic law of large numbers for mean field interacting particles, see a nice informal discussion in [16], where also other popular types of interaction geometry (e.g. d-dimensional lattice and the hierarchical interactions) are touched upon.

Another special class of interactions, namely k-ary group interaction of arbitrary order k (possibly non-binary, and possibly non preserving the total number of particles), was addressed in [7] (developing further some ideas from [6]), where it was shown that if the k-ary interactions of indistinguishable particles that preserve the number of particles are described by generators (conditionally positive operators) $\mathcal{B}_k$ in $C_\infty(X^k)$, $k = 1, \ldots, K$, and the k-ary interactions that change the number of particles (from k to m) are given by symmetrical transitional kernels $P_k(x_1, \ldots, x_k; dy_1 \cdots dy_m)$, the formal measure-valued limit of the large number of particles is given by the equation

$$\frac{d}{dt}(g,\mu_t) = \sum_{k=1}^{K} \frac{1}{k!} \int [(\mathcal{B}_k g^+)(x_1,\dots,x_k)$$
$$+ \sum_m \int (g^+(y_1,\dots,y_m) - g^+(x_1,\dots,x_k)) P_k(x_1,\dots,x_k; dy_1\cdots dy_m)] \mu_t(dx_1)\cdots\mu_t(dx_k), \tag{8}$$

where $g^+(x_1,\dots,x_l) = g(x_1)+\cdots+g(x_l)$ (the weak form (8) being introduced in [39], see also Sections 5 and 7 for details). Notice that though the original interacting particle systems leading to (5), (7) (mean field interaction preserving the number of particles) and to (8) (including k-ary jumps with possible fragmentation or coagulation) are quite different, equation (8) can be written in form (5). Moreover, it is easy to understand that equation (8) describing the limit of systems with interactions not preserving the number of particles, can be also written in form (6). In fact, if all $m \geq k$, then (due to the symmetry of transition kernels)

$$\int g^+(y_1,\dots,y_m) P_k(x_1,\dots,x_k; dy_1\cdots dy_m)$$
$$= \int \frac{m}{k} g^+(y_1,\dots,y_k) P_k(x_1,\dots,x_k; dy_1\cdots dy_m)$$
$$= m \int g^+(y_1,\dots,y_k) \tilde{P}_k^m(x_1,\dots,x_k; dy_1\cdots dy_k)$$

with

$$\tilde{P}_k^m(x_1,\dots,x_k; dy_1\cdots dy_k) = \frac{1}{k}\int_{X^{m-k}} P_k(x_1,\dots,x_k; dy_1\cdots dy_m)$$

and introducing the operators

$$B_k f(x_1,\dots,x_k) = \mathcal{B}_k f(x_1,\dots,x_k)$$
$$+ \sum_m \int [m(f(y_1,\dots,y_k) - kf(x_1,\dots,x_k)] \tilde{P}_k^m(x_1,\dots,x_k; dy_1\cdots dy_k)) \tag{9}$$

one can rewrite (8) as

$$\frac{d}{dt}(g,\mu_t) = \sum_{k=1}^{K} \frac{1}{k!} \int (B_k g^+)(x_1,\dots,x_k) \mu_t(dx_1)\cdots\mu_t(dx_k) \tag{10}$$

with conditionally positive operators B_k in $C_\infty(X^k)$. By similar manipulations (that we omit) one can deal with the case $m < k$.

Equation (10) has the form (6) with

$$A_k g = \frac{1}{k!} B_k g^+. \tag{11}$$

Characterizing positivity preserving evolutions in case of X being a Euclidean space (or a manifold) we shall firstly specify the class of polynomial equations (6) that can be presented in form (10) with conditionally positive operators B_k, then give a structure of arbitrary equations (6) preserving positivity, and at last show (under mild technical assumptions) that a nonlinear operator A in positivity preserving equations (5) has to be of form (7). It will turn out, in particular, that in case $K = 2$ the positivity preserving equations (6) can be always written in form (10) with conditionally positive operators B_k (which is not the case neither for a discrete X, nor for the Euclidean X and $K > 2$).

1.3 Content of the paper and bibliographical comments

At the end of this introduction we are going first to illustrate the difference between positivity preserving evolutions (6) and kinetic equations (10) (in the terminology of the next section this means the difference between conditionally positive and strongly conditionally positive operators) on a simple case of a discrete state space X. At last we shall fix some basic notations to be used in the paper without further reminder.

In Section 2 our main structural results on the positivity preserving generators A, A_k in (5), (6) are obtained (Theorems 2.2-2.4) for the case of X being a Euclidean space (or similarly a manifold). Some examples are given. At the end of Section 2 we discuss shortly the lifting of the nonlinear evolutions (6) to linear evolutions on the multi-particle state space.

As an important previous contribution to the theory of positivity preserving equations of type (4) one should mention book [54], where these equations were analyzed by introducing a natural infinite-dimensional smooth manifold structure on the space of measures and its tangent spaces and where one can find a result, which is essentially equivalent to our Theorem 2.4.

As in the case of linear generators, one thing is to give a structure of formal generators, and another problem is to distinguish those of these formal generators that actually generate semigroups. Though lots of work in this direction is done both for linear and nonlinear cases, many questions remain open. Recent achievements in the linear theory are nicely presented in [28]. There exists an extensive literature on the well posedness for equation (5), (7) with differential generators A that appear in the analysis of interacting diffusions (second order operators A), see [49], [55], [56], [24], [59] and references therein, and of the deterministic processes (first order operators A), see e.g. [10], [19]. Integral operators A, A_k in (6), (7) stand for interacting pure jump processes and their study was motivated by the Boltzmann model of collisions and the Smoluchovski model of coagulation and fragmentation, see e.g. [44], [52], [46], [38], [39], [42] and references therein. Some particular combinations of differential and integral generators are mostly centered around spatially nontrivial versions of Boltzmann and Smoluchovski models, see [25], [27], [50],[35] and references therein, see also [23] for models with spontaneous births.

In recent years one observes a new wave of interest to the Lévy processes (see e.g. books [3], [4]) and their spatially nontrivial extensions - Lévy type processes, the latter being basically just the Markov processes with pseudo-differential generators of Lévy-Khinchin type, see [28] for a comprehensive analytic study of these processes, [29] for a probabilistic analysis, where these processes are called jump-diffusions, and also [53] for some connections between these approaches. Lévy and Lévy type processes enjoy quite different properties (e.g. fat tails and discontinuous trajectories) as usual Wiener process and related diffusions, which make Lévy processes indispensable for applications in many situations where diffusion processes turn out to be un-adequate. On the other hand, the analysis of these processes is surely more complicated than that of usual diffusions, due in particularly to technically more involved analysis of the corresponding stochastic differential equations (see e.g. [3] and [5] for the latter). Roughly speaking, the results of our Section 2 state that positivity preserving measure valued evolutions under mild technical assumptions have generators of non-linear spatially nontrivial Lévy-Khinchin type thus representing nonlinear analogs of Lévy type processes and semigroups.

In Sections 3 we start our analysis of general positivity preserving evolutions by establishing (using a fixed point argument that is well known in nonlinear analysis, see e.g. [48] for its application to nonlinear Schrodinger equation and [55] for stochastic Ito's equations describing the evolution (5) with differential generators) a general well posedness result (Theorem 3.1) allowing to reduce the study of nonlinear problems (5)-(7) to some regularity properties of the corresponding linear problems. This result can be applied to a variety of models with pseudo-differential generators (beyond standard diffusions), where the corresponding regularity results for linear generators are available, for instance in case of decomposable generators (see [37]), regular enough degenerate diffusions like interacting curvilinear Ornstein-Uhlenbeck processes or stochastic geodesic flows on the cotangent bundles to Riemannian manifolds (see [1], [33]), or stable jump diffusions. As the latter are of special interest among Lévy type processes, due to their applications in a wide variety of models from plasma physics to finances (see e.g. [3], [8], [4], [58]), we shall concentrate further on stable type processes ending Section 3 with a simple illustration of Theorem 3.1 in the case of non-degenerate second order part of a pseudo-differential generator.

In Section 4 we introduce our basic model of interacting stable jump-diffusions, prove a well posedness result (reducing exposition for simplicity to the case where the stable part is not affected by interaction) and analyze the core of the generator of the corresponding (linear) Feller process in $C(\mathcal{M}^+(X))$.

In Sections 5 and 6 the general analytic tools for proving the law of large numbers and central limit theorems respectively are developed for interacting Markov processes with pseudo-differential generators. These tools are illustrated on our basic example - interacting stable jump-diffusions, but they seem to be of interest even for usual interacting diffusions. Unlike a probabilistic approach (compactness in Skorohod space of trajectories) normally used for obtaining these type of results (see e.g. [24], [27], [52], [38] and references therein for the law of large numbers, [13], [17], [57], [55], [56], [22] and references therein for the central limit of interacting diffusions and [18], [20], [23], [50] for the central limit to some modes with jumps), our method, being based exclusively on the theory of contraction semigroups, yields new insights in the analytic aspects of the processes under consideration (Feller property, structure of the cores of the generators, precise rate of convergence, etc). Moreover, instead of describing the limiting Gaussian process of fluctuations in the very large space $S'(X)$, as is usual in the literature, our method allows to do it in a natural smaller space$(C^1(X))^*$ that is just "a bit larger", than the initial space $\mathcal{M}(X)$ (which itself is usually unappropriate for describing the fluctuations). Another specific technical feature is the systematic use of the derivatives of the solutions of the kinetic equations with respect to the initial data (in case of the Botzmann equation these derivatives being studied in detail in [40]).

Section 7 is devoted to the interaction changing the number of particles, whose law of large number is described by equation (8). As we noted above this equation can be also written in form (10) with conditionally positive operators B_k in $C(X^k)$, which allows to consider them as particular cases of (10). However, for the application to interacting particles an alternative reduction to (10) becomes more natural, namely by observing that (8) has already form (10) but with B_k presenting not just the operators in $C^{sym}(X^k)$ (as we always supposed in (10)), but the operators from $C^{sym}(\mathcal{X})$ to $C^{sym}(X^k)$ given by

$$B_k f(x_1, \dots, x_k) = \mathcal{B} f(x_1, \dots, x_k) + \sum_{m=1}^{\infty} \int (f(y_1, \dots, y_m) - f(x_1, \dots, x_k)) P_k(x_1, \dots, x_k; dy_1 \cdots dy_m) \qquad (12)$$

that are conditionally positive in the sense that if a non-negative $f \in C^{sym}(\mathcal{X})$ vanishes at $\mathbf{x} = (x_1, \dots, x_k)$, then $B_k f(x) \geq 0$. The theory of Sections 5 and 6 is presented in a way that more or less straightforwardly extends to the Markov models of interactions not preserving the number of particles thus including the processes of, say, coagulation and fragmentation. In Section 7 we illustrate this development by obtaining new results on the rate of convergence to the law of large numbers and the central limit in the standard Smoluchovski model of coagulation-fragmentation (thus giving a solution to Problem 10 from [2] for the case of bounded intensities).

In Appendix two simple auxiliary results are presented, the first of which is devoted to the possibility of approximation of continuous functions on measures by analytic functions in such a way that this approximation respects the derivation (that does not follow from the standard Stone-Weierstrass theorem on approximation).

It is worth noting that the obtained limits of fluctuation processes supply a large class of natural examples for infinite dimensional Mehler semigroups, whose analysis is well under way in the present literature, see e.g. [45] and references therein for general theory, [51] for some properties of Gaussian Mehler semigroups and [15] for the connection with branching processes with immigration.

1.4 Remarks on discrete case

As a warming up we shall discuss shortly the case of X being the set of natural numbers and the r.h.s of (6) being a homogeneous quadratic polynomial. In this case equation (6) takes the form of the system of quadratic equations

$$\dot{x}_j = (A^j x, x), \qquad j = 1, 2, \dots, N, \qquad (13)$$

where N is a natural number or $N = \infty$, unknown $x = (x_1, x_2, \dots)$ is N-vector and all A^j are given square $N \times N$-matrices. If X is the set of natural numbers and only binary interactions preserving the number of particles are allowed, measure-valued kinetic equation (8) is reduced to the discrete system

$$\dot{x}_j = \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{K} x_k x_l \sum_{n \leq m} P_{kl}^{nm} (\delta_n^j + \delta_m^j - \delta_k^j - \delta_l^j), \qquad j = 1, 2, \dots, N, \qquad (14)$$

where P_{kl}^{nm} is an arbitrary collection of positive numbers that is symmetric with respect to the exchange of n and m (defining the rates of transformation of any pair of particles j of type k and l to the pair of particles of type m and n), and where δ_n^j is the Kronecker symbol (equals 1 if $j = n$ and 0 otherwise).

Proposition 1.1 *Suppose N is finite and $\sum_{j=1}^{N} A^j = 0$ (this assumption is made in order to avoid additional problems with the existence of the solutions).*

(i) System (13) *defines a positivity preserving semigroup (i.e. if all coordinates of the initial vector* x_0 *are non-negative, then the solution* $x(t)$ *is globally uniquely defined and all coordinates of this solution are non-negative for al times), if and only if for each* j *the matrix* $\tilde{A}^j$ *obtained from* A^j *by deleting its* j*-th column and* j*-th row is such that* $(\tilde{A}^j v, v) \geq 0$ *for any* $(N-1)$*-vector* v *with non-negative coordinates.*

(ii) System (13) *can be written in form* (14) *(with some positive numbers* P_{kl}^{nm} *) if and only if the entries* A_{kl}^j *are non-negative whenever* $k \neq j$, $l \neq j$ *, i.e. if and only if the matrix* $\tilde{A}^j$ *has only non-negative entries for all* j *.*

We omit a simple proof of this fact. It was presented to give a feeling of the difference between the class of positivity preserving evolutions and the class of kinetic equations (10). This difference will be studied systematically in the next section.

1.5 Notations for spaces and semigroups

Throughout the paper the letter X denotes a locally compact metric space, and $C(X)$ (respectively $C_\infty(X)$) denotes the Banach space of bounded continuous functions on X (respectively its closed subspace consisting of functions vanishing at infinity) equipped with the norm $\|f\|_{C(X)} = \sup_{x\in X} |f(x)|$. Denoting by X^0 a one-point space and by X^j the powers $X \times \cdots \times X$ (j-times), we shall denote by $\mathcal{X}$ their disjoint union $\mathcal{X} = \bigcup_{j=0}^\infty X^j$, which is again a locally compact space. The elements of $\mathcal{X}$ will be designated by bold letters, e.g. $\mathbf{x}, \mathbf{y}$. Instead of $\mathcal{X}$ it is often more convenient to work with its symmetrization $S\mathcal{X}$ which is the quotient space obtained by identifying any two pairs $\mathbf{x}, \mathbf{y}$ that differ only by a permutation of its elements. We shall denote by $C^{sym}(\mathcal{X}) = C(S\mathcal{X})$ (resp. $B^{sym}(\mathcal{X})$) the Banach spaces of symmetric continuous (respectively bounded measurable) functions on $\mathcal{X}$ and by $C^{sym}(X^k)$ (resp. $B^{sym}(X^k)$) the corresponding spaces of functions on the finite powers X^k . $\mathcal{M}(X)$ (respectively $\mathcal{M}_{sym}(\mathcal{X})$) denotes the Banach space of finite Borel measures on X (respectively symmetric measures on $\mathcal{X}$) with the full variation as the norm. By weak topology in the space of measures we shall always understand the $*$-weak topology. The upper subscript "+" for all these spaces (e.g. $\mathcal{M}^+(X)$) will denote the corresponding cones of non-negative elements. The lower subscript "fin" for the spaces of functions on $\mathcal{X}$, e.g. $C^{sym}_{fin}(\mathcal{X})$, will denote a subspace with only finite number of non-vanishing components (polynomial functionals).

We shall work mostly with $X = \mathbf{R}^d$ or X being a closed manifold. In this case by $C^k(X)$ we shall denote the Banach space of bounded k times continuously differentiable functions with bounded derivatives equipped with the norm

$$\|f\|_{C^k(X)} = \sum_{l=0}^k \|f^{(l)}\|_{C(X)}.$$

Of course $C(X) = C^0(X)$ and $C^k_c(X)$ (respectively $C^k_\infty(X)$) denotes the subspace of $C^k(X)$ consisting of functions with a compact support (respectively vanishing at infinity together with all their derivatives up to and including the order k).

By $L_1(X)$ we denote the usual Banach space of integrable functions, and $W^k(X)$ is the Sobolev space of measurable functions having finite norm

$$\|f\|_{W^k(X)} = \sum_{t=0}^k \|f^{(t)}\|_{L_1(X)},$$

where the derivatives are defined in the weak sense. At last, for a Banach space B and a positive number t_0 , we shall denote by $C([0,t_0],B)$ the Banach space of continuous functions from $[0,t_0]$ to B equipped with the norm

$$\|f_t\|_{C([0,t_0],B)} = \sup_{t\in[0,t_0]} \|f_t\|_B.$$

By a propagator (respectively a backward propagator) in a family of sets M_s , $s \geq 0$, we shall mean a family $U(t,s)$, $t \geq s \geq 0$ of transformations $M_s \mapsto M_t$ (respectively $M_t \mapsto M_s$) satisfying the co-cycle identity $U(t,s)U(s,r) = U(t,r)$ (respectively $U(s,r)U(t,s) = U(t,r)$) for $t \geq s \geq r$ and such that $U(t,t)$ is the identity operator for any t. In particular, *a propagator of probability* in a family of the Borel subsets X_t of X (shortly, in X) is a propagator of linear contractions $V(t,s) : \mathcal{M}^+(X_s) \mapsto \mathcal{M}^+(X_t)$. The adjoint operators $U(t,s) = V^*(t,s)$ act backwards from $B(X_t)$ to $B(X_s)$ (and form a backward propagator), the duality equation being

$$(U(t,s)f_t, \mu_s) = (f_t, V(t,s)\mu_s)$$

for functions $f_t \in B(X_t)$ and measures $\mu_s \in \mathcal{M}^+(X_s)$. We say that such a propagator is Feller if the family $U(t,s)$ is a strongly continuous family of operators $C_\infty(X_t) \mapsto C_\infty(X_s)$ (if $U(t,s)$ depend only on the difference $t-s$, these operators form a Feller semigroup). A Markov process specified by $U(t,s)$, i.e. by the transition probabilities $P(t,s,x,dy)$ from X_s to X_t defined by

$$\mathbf{E}_x^{t,s} f_t = \int f_t(y) P(t,s,x,dy) = (U(t,s)f_t)(x),$$

will be then called a nonhomogeneous Feller process (where we use the standard notations $\mathbf{E}_x$ for the expectation of the process starting at x).

2 The structure of generators

Definition. Suppose D is a linear hull of a dense subspace in $C_\infty(X)$ and the set of constant functions. A mapping Ω from a dense subspace of $\mathcal{M}(X)$ containing finite linear combinations of Dirac measures to (possibly unbounded) linear functionals in $C(X)$ with domains containing D is called $\mathbf{x} = (x_1, \dots, x_n)$-*conditionally positive* for a given finite collection of elements $\mathbf{x}$ of X, if $\Omega(\sum_{i=1}^n \omega_i \delta_{x_i})g$ is non-negative for arbitrary positive numbers $\omega_1, \dots, \omega_n$ whenever $g \in D$ is non-negative and vanishes at points $x_1, \dots, x_n$. Ω is called *conditionally positive* if it is $\mathbf{x}$-*conditionally positive* for all $\mathbf{x} \in \mathcal{X}$.

In particular, applying this definition to polynomial operators Ω (see the r.h.s. of equation (6)) we say that a linear map

$$A = (A_1, \dots, A_K) : D \mapsto (C(X), C^{sym}(X2), \dots, C^{sym}(X^K)) \tag{15}$$

is *conditionally positive* if for any collection of different points $x_1, \dots, x_m$ of X , for any non-negative function $g \in D$ such that $g(x_j) = 0$ for all $j = 1, \dots, m$, and for any collection of positive numbers ω_j one has

$$\sum_{k=1}^K \sum_{i_1=1}^m \cdots \sum_{i_k=1}^m \omega_{i_1} \cdots \omega_{i_k} A_k g(x_{i_1}, \dots x_{i_k}) \geq 0. \tag{16}$$

Moreover, we shall say that A is *strongly conditionally positive* if $A_k g(x_1, \dots, x_k)$ is $\mathbf{x} = (x_1, \dots, x_k)$-conditionally positive for any k and $\mathbf{x}$. Finally, a linear operator $A_k : D \mapsto C^{sym}(X^k)$ is said to be *strongly conditionally positive* or *conditionally positive* if this is the case for the family $A = (0, \dots, 0, A_k)$ of type (15).

Remark. If Ω is linear, i.e. $\Omega(\mu)g = (Ag, \mu)$ with some linear operator A in $C(X)$, the above definition of conditional positivity is reduced to the standard definition of conditional positivity of the linear operator A (see introduction).

Our definitions are motivated by the following simple observations.

Proposition 2.1 *(i) If the solutions to* (4) *are defined at least locally and are positive for initial measures being positive finite linear combinations of Dirac delta-measures so that* (4) *holds for all* $g \in D$*, then* Ω *is conditionally positive.*

(ii) Suppose linear operators B_k *in* $C_\infty(X^k)$*,* $k = 1, \dots, K$ *, are conditionally positive in the usual sense (and thus specify a kinetic equation* (10)*). Then the operators* $g \mapsto B_k g^+$ *from* $C_\infty(X)$*(dense subspace of) to* $C_\infty(X^k)$ *are strongly conditionally positive (the notation* g^+ *is introduced in* (8)*).*

Proof. (i) Let a non-negative $g \in D$ be such that $g(x_1) = \cdots = g(x_m) = 0$ for some points $x_1, \dots, x_m$. Choosing $\mu_0 = \omega_1 \delta_{x_1} + \cdots + \omega_m \delta_{x_m}$ we see that $(g, \mu_0) = 0$, and consequently the condition of positivity preservation implies that $\frac{d}{dt}(g, \mu_t) \mid_{t=0} \geq 0$, which means that $\Omega(\sum_{i=1}^n \omega_i \delta_{x_i})g$ is non-negative. Statement (ii) is obvious.

Remark. Thus on the formal level, the difference between the general positivity preserving evolutions of form (6) and the kinetic equations of k-ary interacting particles (10) is the difference between conditionally positive and strongly conditionally positive generators. For the case of discrete X this observation is illustrated by Proposition 1.1.

Our first structural result characterizes the conditional positivity at fixed points.

Theorem 2.1 *Suppose* $X = \mathbf{R}^d$ *with a natural* d *and the space* D *from the definition above contains* $C_c^2(X)$*. If an operator is* $\mathbf{x} = (x_1, \dots, x_n)$*-conditionally positive, then*

$$\Omega\left(\sum_{i=1}^n \omega_i \delta_{x_i}\right) g = \sum_{j=1}^n \Big[a^j(\mathbf{x}) g(x_j) + (b^j(\mathbf{x}), \nabla) g(x_j) + \frac{1}{2} (c^j(\mathbf{x}) \nabla, \nabla) g(x_j)$$

$$+ \int (g(x_j + y) - g(x_j) - \chi(y)(y, \nabla) g(x_j)) \nu^j(\mathbf{x}; dy) \Big] \qquad (17)$$

for $g \in C_c^2(X)$*, where each* c^j *is a positive definite matrix, each* ν^j *is a Lévy measure,* χ *is an indicator like in* (1)*, and with all* a^j, b^j, c^j, ν^j *depending on* $\sum_{i=1}^n \omega_i \delta_{x_i}$ *.*

Proof. Let us start with a comment on the Courrège theorem. A look on the proof of his theorem (see [12], [9], [28]) shows that the characterization is actually given not only for conditionally positive operators, but also for conditionally positive linear functionals obtained by fixing the arguments. Namely, it is shown that if the linear functional $(Ag)(x) : C_c^2 \mapsto \mathbf{R}^d$ is *conditionally positive* at x, i.e. if $Ag(x) \geq 0$ whenever a non-negative g vanishes at x, then $Ag(x)$ has form (1) irrespectively of the properties of $Ag(y)$ in other points y . Now let us choose a partition of unity as a family of n smooth non-negative functions χ_i, $i = 1, \dots, n$, such that $\sum_{i=1}^n \chi_i = 1$ and each χ_i equals one in a neighborhood of x_i (and consequently vanishes in a neighborhood of any other point

x_l with $l \neq i$). By linearity $\Omega(\sum_{i=1}^n \omega_i \delta_{x_i}) = \sum_{i=1}^n \Omega_i$, with $\Omega_j g = (\sum_{i=1}^n \omega_i \delta_{x_i})(\chi_j g)$.. Clearly each functional $\Omega_i g$ is conditionally positive at x_i in the usual sense, i.e. $g(x_i) = 0$ implies $\Omega_i g \geq 0$ for a non negative function $g \in C_c^2(X)$. Hence applying a "fixed point version" of the Courrège theorem to each Ω_i, one obtains representation (17).

As a corollary we shall obtain now a characterization of the strong conditional positivity.

Theorem 2.2 *Let X and D be the same as in Theorem 2.1.*

(i) A linear operator $A_k : D \mapsto C^{sym}(X^k)$ is strongly conditionally positive if and only if

$$A_k g(x_1, \ldots, x_k) = \sum_{j=1}^k [a_k \Pi_{j1}(x_1, \ldots, x_k))g(x_j)$$
$$+ (b_k \Pi_{j1}(x_1, \ldots, x_k)), \nabla)g(x_j) + \frac{1}{2}(c_k \Pi_{j1}(x_1, \ldots, x_k))\nabla, \nabla)g(x_j)$$
$$+ \int (g(x_j + y)g(x_i) - \chi(y)(y, \nabla)g(x_j))\nu_k(\Pi_{j1}(x_1, \ldots, x_k); dy)], \qquad (18)$$

where the operators Π_{jl} in $\mathbf{R}^d$ exchange the coordinates with numbers j and l, each c_k (respectively ν^j) is a positive definite matrix (respectively a Lévy measure), χ is an indicator like in (1)*, and where the functions a_k, b_k, c_k, ν_k are symmetric with respect to the permutations of the variables $x_2, \ldots, x_k$ (i.e. all permutations not affecting x_1).*

(ii) A linear operator $A_k : D \mapsto C^{sym}(X^k)$ is strongly conditionally positive if and only if

$$A_k g(x_1, \ldots, x_k) = \sum_{j=1}^k a_k(\Pi_{j1}(x_1, \ldots, x_k))g(x_j) + B_k g^+(x_1, \ldots, x_k), \qquad (19)$$

where a_k is the same as in (17) *and B_k is a conditionally positive (in usual sense) operator in $C_\infty^{sym}(X^k)$.*

(iii) From the continuity of the functions in the image of A it follows that the functions a_k, b_k, c_k, ν_k depend continuously on $x_1, \ldots, x_k$ (measures ν_k are considered in the weak topology).

Proof. (i) From Theorem 2.1 it follows that for arbitrary different $x_1, \ldots, x_k$

$$A_k g(x_1, \ldots, x_k) = \sum_{j=1}^k [a_k^j(x_1, \ldots, x_k)g(x_j)$$
$$+ (b_k^j(x_1, \ldots, x_k), \nabla)g(x_j) + \frac{1}{2}(c_k^j(x_1, \ldots, x_k)\nabla, \nabla,)g(x_j)$$
$$+ \int (g(x_j + y) - g(x_j) - \chi(y)(y, \nabla)g(x_j))\nu_k^j(x_1, \ldots, x_k; dy)],$$

where each c_k^j is a positive definite matrix, each ν_k^j is a Lévy measure. Moreover, due to the symmetry of functions $A_k g$ for $g \in C_c^2(X)$, the functions $a_k^j, b_k^j, c_k^j, \nu_k^j$ depend on $x_1, \ldots, x_k$ symmetrically in the sense that, say $a^j \Pi_{jl} = a^l$ for all j, l and $a^j \Pi_{lm} = a^j$ for

$m \neq j$, $l \neq j$. Due to this symmetry it is possible to rewrite the above formula as (18) 1 with $a_k = a_k^1$, $b_k = b_k^1$, $c_k = c_k^1$, $\nu_k = \nu_k^1$. It remains to notice that due to the continuity, this representation remains valid even for not necessarily different points $x_1, \dots, x_k$.

(ii) By the Courrège theorem applied to X^k , a conditionally positive B_k in $C_\infty(X^k)$ has form

$$B_k f(x_1, \dots, x_k) = \tilde{a}(x_1, \dots, x_k) f(x_1, \dots, x_k)$$
$$+ \sum_{j=1}^{k} (\tilde{b}^j(x_1, \dots, x_k), \nabla_{x_j}) f(x_1, \dots, x_k)) + \frac{1}{2}(\tilde{c}(x_1, \dots, x_k)\nabla, \nabla) f(x_1, \dots, x_k)$$
$$+ \int (f(x_1 + y_1, \dots, x_k + y_k) - f(x_1, \dots, x_k) - \sum_{i=1}^{k} (\chi(y_i)(y_i, \nabla_{x_i}) f(x_1, \dots, x_k))$$
$$\times \tilde{\nu}(x_1, \dots, x_k; dy_1 \cdots dy_k). \tag{20}$$

Applying this to $f = g^+$ and comparing with (17) yields the required result.

(iii) By one and the same procedure, one proves first the continuity for different $x_1, \dots, x_k$, then for two of them coinciding, etc. If all $x_1, \dots, x_k$ are different, one can use the decomposition from the proof of Theorem 2.1 reducing the problem to proving the continuity for the operator $A_k(\chi_j g)(x_1, \dots, x_k)$. Choosing g to be a constant, one proves the continuity of the coefficients a_k for this operator. Then choosing g to be a constant in a neighborhood of a certain point x one proves the continuity of ν_k . Then choosing g to be linear around x one gets the continuity of b_k , and at last the continuity of c_k follows.

Thus operators A_k in (4) are strongly conditionally positive if and only if they have form $A_k g = B_k g^+/k!$ with some conditionally positive B_k in $C_\infty(X)$ (up to some multiplication operators) and thus if they correspond to a kinetic equation for some Markov model of k-ary interacting particles.

Theorem 2.3 *Suppose again that $X = \mathbf{R}^d$ and D contains $C_c^2(X)$. A linear mapping* (15) *is conditionally positive if and only if each of the operators A_k has form* (18) *with the same symmetry condition on its coefficients as in* (18) *and with c_k and ν_k being not necessarily positive but only such that the matrix*

$$\sum_{k=1}^{K} k \sum_{i_1=1}^{m} \cdots \sum_{i_{k-1}=1}^{m} \omega_{i_1} \cdots \omega_{i_{k-1}} c_k(x, x_{i_1}, \dots, x_{i_{k-1}}) \tag{21}$$

is positive definite and the measure

$$\sum_{k=1}^{K} k \sum_{i_1=1}^{m} \cdots \sum_{i_{k-1}=1}^{m} \omega_{i_1} \cdots \omega_{i_{k-1}} \nu_k(x, x_{i_1}, \dots, x_{i_{k-1}}) \tag{22}$$

is positive for any m, any collections of positive numbers $\omega_1, \dots, \omega_m$ and points $x, x_1, \dots, x_m$ (with all ν_k satisfying (2)*, of course). In particular, A_k have form* (19) *with B_k given by* (20) *but with not necessarily positive $\tilde{c}$ and $\tilde{\nu}$.*

Proof. **Step 1**. Here we shall prove that all A_k have form (17) with $n = k$, but possibly without positivity of c_k and ν_k . For arbitrary fixed $x_1, \dots, x_m$ the functional of g given by

$$\sum_{k=1}^{K} \sum_{i_1=1}^{m} \cdots \sum_{i_k=1}^{m} \omega_{i_1} \cdots \omega_{i_k} A_k g(x_{i_1}, \dots, x_{i_k}) \tag{23}$$

is $\mathbf{x} = (x_1, \dots, x_m)$- conditionally positive and consequently has form (17) (with m instead of n), as follows from Theorem 2.1. Using $\epsilon\omega_j$ instead of ω_j, dividing by ϵ and then letting $\epsilon \to 0$ one obtains that $\sum_{i=1}^m A_1 g(x_i)$ has the same form. As m is arbitrary, this implies on the one hand that $A_1 g(x)$ has the same form for arbitrary x (thus giving the required structural result for A_1) and on the other hand that

$$\sum_{k=2}^K \epsilon^{k-2} \sum_{i_1=1}^m \cdots \sum_{i_k=1}^m \omega_{i_1} \cdots \omega_{i_k} A_k g(x_{i_1}, \dots, x_{i_k})$$

has the required form. Again letting $\epsilon \to 0$ yields the same representation for the functional

$$\sum_{i_1=1}^m \sum_{i_2=1}^m \omega_{i_1} \omega_{i_2} A_2 g(x_{i_1}, x_{i_2}).$$

As above this implies on the one hand that

$$\omega_1^2 A_2 g(x_1, x_1) + 2\omega_1\omega_2 A_2 g(x_1, x_2) + \omega_2^2 A_2 g(x_2, x_2)$$

has the required form and hence also $A_2 g(x_1, x_1)$ has this form for arbitrary x_1 (put $\omega_2 = 0$ in the previous expression), and hence also $A_2 g(x_1, x_2)$ has the same form for arbitrary x_1, x_2 (thus giving the required structural result for A_2), and on the other hand that

$$\sum_{k=3}^K \epsilon^{k-3} \sum_{i_1=1}^m \cdots \sum_{i_k=1}^m \omega_{i_1} \cdots \omega_{i_k} A_k g(x_{i_1}, \dots, x_{i_k})$$

has the required form. Following this procedure inductively yields the claim of Step 1.

Step 2. From the obtained representation for A_k it follows that

$$\sum_{k=1}^K \sum_{i_1=1}^m \cdots \sum_{i_k=1}^m \omega_{i_1} \cdots \omega_{i_k} A_k g(x_{i_1}, \dots, x_{i_k}) = \sum_{k=1}^K k \sum_{l=1}^K \sum_{i_1=1}^m \cdots \sum_{i_{k-1}=1}^m \omega_l \omega_{i_1} \cdots \omega_{i_{k-1}}$$

$$\Big[a_k(x_l, x_{i_1}, \dots, x_{i_{k-1}}) g(x_l) + (b_k(x_l, x_{i_1}, \dots, x_{i_{k-1}}), \nabla) g(x_l) + \frac{1}{2} (c_k(x_l, x_{i_1}, \dots, x_{i_{k-1}}) \nabla, \nabla) g(x_l)$$

$$+ \int (g(x_j + y) - g(x_j) - \chi(y)(y, \nabla) g(x_j)) \nu_k(x_l, x_{i_1}, \dots, x_{i_{k-1}}; dy) \Big].$$

As this functional has to be strongly conditionally positive, the required positivity property of c and ν follows from Theorem 2.1.

Corollary 1 *In the case $K = 2$ the notions of conditional positivity and strong conditional positivity for* (15) *coincide.*

Proof. In case $K = 2$ the positivity of (21) reads as the positivity of the matrix

$$c_1(x) + 2 \sum_{i=1}^m \omega_i c_2(x, x_i) \tag{24}$$

for all natural m, positive numbers ω_j and points x, x_j, $j = 1, \dots, m$. Hence c_1 is always positive (put $\omega_j = 0$ for all j). To prove strong conditional positivity one has to prove that $c(x,y)$ is positive definite for all x,y. But if there exist x,y such that $c(x,y)$ is not positive definite, then by choosing large enough number of points $x_1,\dots,x_m$ near y , one would get a matrix of form (24) that is not positive definite (even for all $\omega_j=1$). This contradiction completes the proof.

Taking into account explicitly the symmetry of generators in (6) (note that in (6) or (8) the use of nonsymmetric generators or their symmetrizations specifies the same equation) allows to get a useful equivalent representation for (5), (7). Namely, the following statement is obvious.

Corollary 2 *(i) If A is conditionally positive and hence by Theorem 2.2 its components have form* (18)*, one can rewrite* (6) *as*

$$\frac{d}{dt}(g,\mu_t)=\sum_{k=1}^K k\int (A_k^1 g)(x,y_1,\dots,y_{k-1})\mu_t(dx)\mu_t(dy_1)\cdots\mu_t(dy_{k-1}), \tag{25}$$

with $A_k^1: C_\infty(X)\mapsto C^{sym}(X^k)$ being defined as

$$A_k^1 g(x,y_1,\dots,y_{k-1})=a_k(x,y_1,\dots,y_{k-1})g(x)$$
$$+(b_k(x,y_1,\dots,y_{k-1}),\nabla g(x))+\frac{1}{2}(c_k(x,y_1,\dots,y_{k-1})\nabla,\nabla)g(x)+\Gamma_k(y_1,\dots,y_{k-1})g(x), \tag{26}$$

where ∇ is, of course, the gradient operator with respect to the variable $x\in R^d$ and where

$$\Gamma_k(y_1,\dots,y_{k-1})g(x)=\int (g(x+z)-g(x)-\chi(z)(z,\nabla)g(x))\nu_k(x,y_1,\dots,y_{k-1};dz). \tag{27}$$

(ii) If A_k are given by (11) *(and hence* (6) *has form* (10)*), then $A_k^1=(1/k!)B_k\pi$, where the lifting operator π is given by $\pi g(x_1,\dots,x_k)=g(x_1)$.*

(iii) The strong form of equation (25)-(27) *for measures μ_t having densities with respect to Lebesgue measure, i.e. having form $\mu_t(dx)=f_t(x)dx$ with some $f_t\in L_1(X)$ is*

$$\frac{d}{dt}f_t(x)=\sum_{k=1}^K k\int \Big[\frac{1}{2}(c_k(x,y_1,\dots,y_{k-1})\nabla,\nabla)f_t(x)+((\nabla c_k-b_k)(x,y_1,\dots,y_{k-1}),\nabla f_t(x))$$
$$+(a_k-\nabla b_k+\frac{1}{2}(\nabla,\nabla c_k))(x,y_1,\dots,y_{k-1})f_t(x)+\Gamma_k^*(y_1,\dots,y_{k-1})f_t(x)\Big]\prod_{l=1}^{k-1}f_t(y_l)dy_l, \tag{28}$$

where

$$(\nabla c_k(x,y_1,\dots,y_{k-1}))^j=\sum_{i=1}^d\frac{\partial}{\partial x_i}c_{ij}(x,y_1,\dots,y_{k-1}),$$
$$(\nabla,\nabla c_k)(x,y_1,\dots,y_{k-1})=\sum_{i,j=1}^d\frac{\partial^2}{\partial x_i\partial x_j}c_{ij}(x,y_1,\dots,y_{k-1}),$$

and where $\Gamma_k^*(y_1,\dots,y_{k-1})$ is the dual operator to the integral part (27) *of* (26).

As a side result, we can give now the structure of positivity preserving evolutions (5).

Theorem 2.4 *Suppose X and D are as in Theorem 2.1. Let A be a mapping from a dense subspace D of $\mathcal{M}(X)$ containing finite combinations of Dirac measures to linear operators in $C_\infty(X)$ with domains containing D. Let the r.h.s. of* (5) *be conditionally positive and let A be continuous in the sense that if $\mu_n \to \mu$ weakly, $\mu, \mu_n \in D$ and $g \in D$, then $A(\mu_n)g \to A(\mu)g$ in $C(X)$. Then $A(\mu)$ has form* (7).

Proof. From Theorem 2.1 and the definition of conditional positivity it follows that $A(\mu)$ has form (7) for μ being finite linear combinations of Dirac measures. For arbitrary $\mu = \lim \mu_n \in D$ with μ_n being finite combinations of Dirac measures, $A(\mu_n)g$ converges to a continuous function for arbitrary $g \in D$. This defines the coefficients a, b, c, ν in (7) as continuous functions for arbitrary $\mu \in \mathcal{D}$ by first proving the continuity of a by choosing $g = 1$, then the continuity of ν by choosing g to be constant in an open set, and then the continuity of b by choosing g to be linear in an open set.

Some remarks and examples to the obtained results are in order.

Remark 1. It is clear that if only two components, say A_i and A_j , do not vanish in the mapping (15), and A is conditionally positive, then each non-vanishing component A_i and A_j is conditionally positive as well (take $\epsilon\omega_j$ instead of ω_j in the definition and then pass to the limits $\epsilon \to 0$ and $\epsilon \to \infty$). In case of more than two non-vanishing components in the family A, the analogous statement is false (in particular, Corollary 1 can not be extended to $K > 2$). Namely, if A is conditionally positive, the "boundary" operators A_1 and A_K are conditionally positive as well (the same argument), but the intermediate operators A_k need not to be, as shows already a simple example of the operator $A = (A_1, A_2, A_3)$ with

$$A_i g(x_1, \dots, x_i)) = a_i(\Delta g(x_1) + \cdots + \Delta g(x_i)), \qquad i = 1, 2, 3,$$

with $a_1 = a_3 = 1$ and with a_2 being a small enough negative number. Of course, it is easy to write an explicit solution to equation (4) in this case.

Remark 2. We gave our results for $X = \mathbf{R}^d$, but using localization arguments (like in linear case, see [9]) the same results can be easily extended to closed manifolds. It is seemingly possible to characterize in the same way the corresponding boundary conditions (generalizing also the linear case from [9]), though this is already not so straightforward.

Remark 3. Basic conditions of positivity of (21), (22) can be written in an alternative integral form. Namely, the positivity of matrices (21) is equivalent (at least for bounded continuous functions c) to the positivity of the matrices

$$\sum_{k=1}^{K} k \int c_k(x, y_1, \dots, y_{k-1}) \mu(dy_1) \cdots \mu(dy_{k-1}) \tag{29}$$

for all non-negative Borel measures $\mu(dy)$. Conditions (21), (22) actually represent modified multi-dimensional matrix-valued or measure-valued versions of the usual notion of positive definite functions. For instance, in case $k = K = 3$ and $d = 1$, (29) means that $\int c_3(x, y, z)\omega(y)\omega(z)dydz$ is a non-negative number for any non-negative integrable function ω. The usual notion of a positive definite function c_3 (as a function of the last two variables) would require the same positivity for arbitrary (not necessarily positive) integrable ω.

Remark 4. As shows Proposition 1.1 the statement of Corollary 1 does not hold for discrete X. In case of continuous state space X one can start feeling the difference between strictly conditionally positive and conditionally positive operators only with $k = 3$. A simple example of a conditionally positive, but not a strictly conditionally positive operator represents the operator defined as

$$A_3 g(x_1, x_2, x_3) = \cos(x_2 - x_3)\Delta g(x_1) + \cos(x_1 - x_3)\Delta g(x_2) + \cos(x_1 - x_2)\Delta g(x_3). \quad (30)$$

Equation (28) in this case takes the form

$$\frac{d}{dt} f_t(x) = \Delta f_t(x) \int f_t(y) f_t(z) \cos(y - z) dydz. \quad (31)$$

The "nonlinear diffusion coefficient" is not strictly positive here. Namely, since

$$\int f(y) f(z) \cos(y - z) dydz = \frac{1}{2}(|\hat{f}(1)|^2 + |\hat{f}(-1)|^2),$$

where $\hat{f}(p)$ is the Fourier transform of f , this expression do not have do be strictly positive for all non-vanishing non-negative f. However, one can find the explicit solution to the Cauchy problem of (31):

$$f_t = \frac{1}{\sqrt{2\pi\omega_t}} \int \exp\left\{-\frac{(x - y)^2}{2\omega_t}\right\} f_0(y) dy,$$

with

$$\omega_t = ln(1 + t(|\hat{f}_0(1)|^2 + |\hat{f}_0(-1)|^2).$$

This is easily obtained by passing to the Fourier transform of equation (31) that has the form

$$\frac{d}{dt} \hat{f}_t(p) = -\frac{1}{2} p^2 (|\hat{f}_t(1)|^2 + |\hat{f}_t(-1)|^2) \hat{f}_t(p),$$

and which is solved by observing that $\xi_t = |\hat{f}_t(1)|^2 + |\hat{f}_t(-1)|^2$ solves the equation $\dot{\xi}_t = -\xi_t^2$ and consequently equals $\xi_t = (t + \xi_0^{-1})^{-1}$.

Remark 5. There is a natural "decomposable" class of operators, for which (6) reduces straightforwardly to a linear problem. Namely, suppose $k! A_k g = B_k g^+ = (\tilde{B}_k g)^+$ for all $k = 1, \dots, K$ with some $\tilde{B}_k$ in $C_\infty(X)$ that generate Markov processes their (i.e. $\tilde{B}_k 1 = 0$). Then (25) takes the form

$$\frac{d}{dt}(g, \mu_t) = \sum_{k=1}^{K} \frac{1}{(k-1)!} (\tilde{B}_k g, \mu_t) \|\mu_t\|^{(k-1)},$$

which is a linear equation depending on $\|\mu_t\| = \|\mu_0\|$ as on a parameter.

For conclusion of this section let us discuss the lifting of (6), (10) to linear evolutions on the multi-particle state space $\mathcal{X}$ which is of importance for the corresponding particle systems describing their "propagation of chaos" property. In fact equations (34), (35) below are derived in [7] from a scaling limit of the moment measures of interacting particle systems in the spirit of e.g. [30] or [56].

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 $J \backslash I$, by $\mathbf{x}_I$ the collection of the variables

$x_{i_1}, \dots, x_{i_k}$ and by $d\mathbf{x}_I$ the measure $dx_{i_1} \cdots dx_{i_k}$. 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 $f(x) = f(x_1, \dots, x_k) = f^k(x_1, \dots, x_k)$ (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

$$(f, \rho) = \int f(x)\rho(dx) = f^0\rho_0 + \sum_{n=1}^{\infty}(x_1, \dots, x_n)\rho(dx_1 \cdots dx_n),$$
$$f \in C^{sym}(\mathcal{X}), \qquad \rho \in \mathcal{M}(\mathcal{X}), \tag{32}$$

so that $\|\rho\| = (\mathbf{1}, \rho)$ for $\rho \in \mathcal{M}^+(\mathcal{X})$. To an arbitrary $Y(dx) \in \mathcal{M}(X)$ there corresponds a measure $Y^{\tilde{\otimes}} \in M_{sym}(\mathcal{X})$ defined by its components

$$(Y^{\tilde{\otimes}})_n(dx_1 \cdots dx_n) = \frac{1}{n!}Y^{\otimes n}(dx_1 \cdots dx_n) = \frac{1}{n!}Y(dx_1) \cdots Y(dx_n). \tag{33}$$

To each $g \in C^{sym}(\mathcal{X})$ there corresponds a *analytic functional* $(g, Y^{\tilde{\otimes}})$ on $\mathcal{M}(X)$ (also called sometimes the generating functional for g). Such a functional will be called a polynomial on $\mathcal{M}(X)$, if $g \in C^{sym}_{fin}(\mathcal{X})$, i.e. if only a finite number of the components of g do not vanish.

Theorem 2.5 *(i) If μ_t satisfies* (10)*, then $\nu_t = (\mu_t)^{\tilde{\otimes}} \in \mathcal{M}_{sym}(\mathcal{X})$ satisfies the linear equation*

$$\frac{d}{dt}\nu_t^l(dx_1 \cdots dx_l) = \sum_{j=1}^{l}\sum_{k=1}^{K} C^l_{l+k-1} \int_{x_{l+1},\dots,x_{l+k-1}} (B_k^{j,l+1,\dots,l+k-1})^* \nu_t^{k+l-1}(dx_1 \cdots dx_{l+k-1}), \tag{34}$$

where C^l_m are the usual binomial coefficients, B^*_k is the dual to B_k and $(B^I_k)^*\nu_t(dx_1 \cdots dx_m)$ means the action of B^*_k on the variables with indexes from $I \subset \{1, \dots, m\}$.

(ii) If the evolution of t $\nu_t \in M_{sym}(\mathcal{X})$ is specified by (34)*, then the dual evolution on $C^{sym}(X)$ is given by the equation*

$$\dot{g}(x_1, \dots, x_l) = (\mathcal{L}_B g)(x_1, \dots, x_l) = \sum_{I \subset \{1,\dots,l\}} \sum_{j \notin I} (B^{j,I}_{|I|+1} g_{\bar{I}})(x_1, \dots, x_l), \tag{35}$$

where $g_I(x_1, \dots, x_l) = g(x_I)$ and $B_k^{j_1,\dots,j_k}$ means the action of B_k on the variables $j_1, \dots, j_k$. In particular,

$$(\mathcal{L}_B g^1)(x_1, \dots, x_l) = (B_l(g^1)^+)(x_1, \dots, x_l).$$

Proof. Observe that the strong form of (10) is

$$\dot{\mu}_t(dx) = \sum_{k=1}^{K} \frac{1}{(k-1)!} B^*_k(\mu_t \otimes \cdots \otimes \mu_t)(dx dy_1 \cdots dy_{k-1}),$$

which implies (34) by straightforward manipulations. From (34) it follows that

$$\frac{d}{dt}(g,\nu_t)=\sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{I\subset\{1,\dots,l+k-1\},|I|=k-1}\sum_{j\notin I}B_k^{j,I}g_{\bar I}(x_1,\dots,x_{l+k-1})\nu_t(dx_1\cdots dx_{l+k-1})$$

$$=\sum_{m=0}^{\infty}\sum_{I\subset\{1,\dots,m\}}\sum_{j\notin I}B_{|I|+1}^{j,I}g_{\bar I}(x_1,\dots,x_m)\nu_t(dx_1\cdots dx_m),$$

which implies (35).

From the duality between (34) and (35) one concludes that whenever one has the well posedness for the Cauchy problem of equations (35) and (10) for some dense subspaces of initial continuous functions g on $\mathcal{X}$ and initial measures μ_0 on X, one has the duality relation

$$(g_t,\mu_0^{\tilde\otimes})=(g_0,\mu_t^{\tilde\otimes}) \tag{36}$$

implying the invariance of the corresponding space of "analytic" functionals of the type $(g,Y^{\tilde\otimes})$ on $\mathcal{M}(\mathcal{X})$ under the action of the semigroup $T_tF(\mu)=F(\mu_t)$ on $C(\mathcal{M}(X))$. In many cases one can prove that this space provides a core for the generator of this semigroup, with the r.h.s. $\mathcal{L}_B$ of (35) specifying the form of the generator on this core.

3 A well posedness result

Suppose $X=\mathbf{R}^d$. Our strategy of solving (28) will be to look for a fixed point of a mapping from $u_t\in C([0,t_0],L_1(X))$ to the solution $F_t\in C([0,t_0],L_1(X))$ of the linear equation

$$\frac{d}{dt}F_t(x)=\sum_{k=1}^{K}k\int\left[\frac{1}{2}(c_k(x,y_1,\dots y_{k-1})\nabla,\nabla)F_t(x)+((\nabla c_k-b_k)(x,y_1,\dots,y_{k-1}),\nabla F_t(x))\right.$$
$$\left.+(\nabla b_k-\frac{1}{2}(\nabla,\nabla c_k)+a_k)(x,y_1,\dots,y_{k-1})F_t(x)+\Gamma_k^*(y_1,\dots,y_{k-1})F_t(x)\right]\prod_{t=1}^{k-1}u_t(y_l)dy_l. \tag{37}$$

Theorem 3.1 *Suppose the coefficients of* (37) *as functions of x belong to $C^2(X)$ with all bounds being uniform with respect to other variables (derivatives of the Lévy measures in the integral operators Γ_k are taken in the weak sense). Suppose for any t_0 and any $u_t\in S_0C([0,t_0],L_1^+(X))$ (the subset in $C([0,t_0],L_1^+(X))$ of functions having fixed L_1-norm $\|u_0\|$ at all times) the resolving operator (or the propagator) $U_{t,s}([u])$, $0\le s\le t\le t_0$ to the Cauchy problem* (37) *in $L_1(X)$ is well defined. More precisely suppose that $U_{t,s}([u])$ is the family of positivity preserving linear isometries in $L_1(X)$ (to have isometries one usually assumes that all a_k vanish in* (37)*) depending strongly continuously on t,s and such that*

(i) $U_{t,t}([u])$ is the identity operator for all t,

(ii) the semigroup identity $U_{t,s}([u])U_{s,r}([u])=U_{t,r}([u])$ holds for all $t\ge s\ge r$,

(iii) the function $g_t=U_{t,s}([u])g_s$ belongs to $C^2(X)$ and is the unique solution to (37) *with the initial condition $g_s=g$ whenever $g\in C_c^2(X)$,*

(iv) the space $W^2(X)$ is preserved by $U_{t,s}([u])$ and

$$\|U_{s,r}([u])g\|_{W^2(X)} \leq C_2(t_0)\|g\|_{X^2(X)} \tag{38}$$

holds with $C_2(t_0)$ depending only on $\|u_0\|$.

Then for an arbitrary nonnegative $f_0 \in W^2(X) \cap C_c^2(X)$ there exists a unique non negative classical solution $f_t \in W^2(X) \cap C^2(X)$ of (28) *with the initial condition f_0 (i.e. it satisfies* (28) *for all $t \geq 0$ and coincides with f_0 at $t = 0$) and the mapping $f_0 \mapsto f_t$ extends to a strongly continuous semigroup of (nonlinear) operators in $(W^2(X))^+$ that preserve the L_1 -norm of f_0 and yield a solution to the Cauchy problem of* (25).

Proof. We are going to show that in case $u_0 \in W^2(X)$ there exists a unique fixed point of the mapping $u_t \mapsto F_t = U_{t,0}([u])u_0$. First observe that differentiating the semigroup identity for $U_{t,s}([u])$ one gets that

$$\frac{d}{ds}U_{t,s}([u])g = U_{t,s}([u])L(s)g \tag{39}$$

for $g \in W^2(X)$, where $L(s)$ denote the generators (depending on u) at time s (the operators on the r.h.s. of equation (37)).

Let u_0 be fixed and let u^1, u^2 coincide with u_0 at time zero and belong to the sphere $S_0C([0, t_0]), L_1^+(X))$. Due to (39) one has

$$\begin{aligned} F_t([u^1]) - F_t([u^2]) &= \int_0^t \frac{d}{ds}U_{t,s}([u^2])U_{s,0}([u^1])u_0\, ds \\ &= \int_0^t U_{t,s}([u^2])(L_1(s) - L_2(s))U_{s,0}([u^1])u_0\, ds, \end{aligned} \tag{40}$$

where $L_1(s), L_2(s)$ denote the generators at times s of the equation (37) with u_1 and u_2 respectively. Clearly

$$\|(L_1(s) - L_2(s))g\|_{L_1(X)} \leq C_1\|u^1(s) - u^2(s)\|_{L_1(X)}\|g\|_{W^2(X)} \tag{41}$$

where C_1 depends on $\|u_0\|_{L_1(X)}$and the bounds for the derivatives of the coefficients of equation (37). Consequently (38), (40), (41) imply that

$$\sup_{0\leq s\leq t}\|F_s([u^1]) - F_s([u^2])\|_{L_1(X)} \leq tC_1C_2(t_0)\|u_0\|_{W^2(X)} \sup_{0\leq s\leq t}\|u^1(s) - u^2(s)\|_{L_1(X)}. \tag{42}$$

Hence the mapping u_t to F_t in the subset of $S_0C([0, t_0], L_1(X))$ obtained by fixing u_0 is a contraction whenever

$$tC_1C_2(t_0)\|u_0\|_{W^2(X)} < 1. \tag{43}$$

Consequently one obtains a unique solution to (39) on the interval $[0, t_1]$ for any $t_1 \leq t_0$ satisfying (43), and moreover

$$\sup_{0\leq s\leq t_1}\|u_s\|_{W^2(X)} \leq C_2(t_0)\|u_0\|_{W^2(X)}. \tag{44}$$

If $t_1 < t_0$ one can repeat this procedure starting from the time t_1 obtaining a unique continuation of the solution to the interval $[t_1, t_2]$, where

$$(t_2 - t_1)C_1(C_2(t_0))^2\|u_0\|_{W^2(X)} < 1. \tag{45}$$

A key observation now is that as this new solution again solves (37) one still has the estimate (44) with t_1 replaced by t_2 (i.e. C_2 do not have to be replaced by C_2^2 here), which allows to attain the time t_0 by using this procedure p times with p being the minimal integer exceeding

$$t_0 C_1 (C_2(t_0))^2 \|u0\|_{W^2(X)}.$$

This completes the proof of the Theorem.

We are going to illustrate the general result obtained by applying it to the case of equations with a strictly non-degenerate diffusion part such that

$$\sum_{k=1}^K k \int c_k(x, y_1, \dots, y_{k-1}) \mu(dy_1) \cdots \mu(dy_{k-1}) \geq \|\mu\|^m \tag{46}$$

for all x and (non-negative) μ and some fixed real m.

Theorem 3.2 *Suppose all a_k vanish in* (25), (26), *c_k satisfy* (46) *and the measures ν_k in* (27) *have densities with respect to Lebesgue measure:*

$$\nu_k(x, y_1, \dots, y_{k-1}; dz) = \psi_k(x, y_1, \dots, y_{k-1}; z) dz \tag{47}$$

such that

$$\psi_k(x, y_1, \dots, y_{k-1}; z) \leq max \left(\frac{C}{|z|^{\beta_1 + d}}, \frac{C}{1 + |z|^{\beta_2 + d}} \right) \tag{48}$$

for all k, x, z, y_j with some constants $\beta_1 \in (0, 2)$, $\beta_2 > 0$, $C > 0$ and with

$$\int_{|y| \leq 1} \sup_{|z| \leq 1} \left| \frac{\partial \psi}{\partial x}(x + z, y) \right| |y|^2 < \infty, \qquad \int_{|y| \leq 1} \sup_{|z| \leq 1} \left| \frac{\partial^2 \psi}{\partial x^2}(x + z, y) \right| |y|^2 < \infty.$$

Suppose that $b \in C^3(X)$ and $c \in C^4(X)$ with all bounds being uniform with respect to other variables. Then all the conditions (and hence the conclusions) of Theorem 3.1 are satisfied.

Proof. Observe first that the dual Γ_k^* to operator (27) has form

$$\Gamma_k^*(y_1, \dots, y_{k-1}) g(x) = \int [g(x - z) \psi(x - z, y_1, \dots, y_{k-1}; z) - g(x) \psi(x, y_1, \dots, y_{k-1}; z)$$
$$+ (\nabla g(x), z) \chi(z) \psi(x, y_1, \dots, y_{k-1}; z) + g(x) \chi(z) (\nabla_x \psi(x, y_1, \dots, y_{k-1}; z), z)] \, dz$$

which can be written in the form

$$\int (g(x - z) - g(x) - (\nabla g(x), z)) \chi(z)) \psi(x - z; z) \, dz$$
$$+ g(x) \int (\psi(x - z; z) - \psi(x; z) + \chi(z) (\nabla_x \psi(x; z), z)) \, dz$$
$$+ (\nabla g(x), \int z \chi(z) (\psi(x, z) - \psi(x - z, z)) \, dz) \tag{49}$$

(where we omitted the arguments $y_1, \dots, y_{k-1}$ of the function ψ), i.e. it has the same form (47) plus a first order differential operator with bounded coefficients (the latter being due to our assumptions on ψ). The well posedness of the Cauchy problem for equation (37) under the conditions of the Theorem and with bounded ψ was shown first seemingly in [43]. In [32], [33] it was shown that in this case there exists a Green function $G(t,x,x_0;[u])$ of equation (37), which is continuously differentiable in t,x,x_0 for $t>0$ and satisfies the Chapman-Kolmogorov equation. In [31] the same result was obtained for Γ_k being a finite sum of fractional powers of the Laplace operator (stable laws). Both proofs from [32] and [31] can be easily generalized to give the existence of a differentiable Green function under the above conditions on ψ . Moreover, as the same holds for the dual equation and the dual equation preserves constants it follows that the solution of (37) preserves the L_1- norm. This shows everything except (38), but the latter estimate is straightforward. In fact, on the one hand, the preservation of smoothness, in particular (38), is well known for non-degenerate diffusion equations, i.e. for $\psi=0$. On the other hand, the series representation for $G(t,x,x_0;[u])$ obtained in [32] for non-vanishing ψ shows that the integral operator Γ can be treated as a small perturbation that does not affect the validity of (38).

4 Interacting stable jump-diffusions

Our basic model of interacting stable jump-diffusions is defined by $X=\mathbf{R}^d$ and $K=2$ in (6) or (10) connected through (11), and by B_1,B_2 given by

$$(B_1f)(x)=-\sigma(x)|\Delta|^{\alpha/2}f(x),\qquad f\in C_c^2(X), \tag{50}$$

$$\begin{aligned}(B_2f)(x,y)&=V(x,y)\frac{\partial f}{\partial x}(x,y)+V(y,x)\frac{\partial f}{\partial y}(x,y)\\ &+\int(f(x+x_1,y+y_1)-f(x,y))\tilde\psi(x,y;x_1,y_1)dx_1dy_1,\qquad f\in C_c^2(X^2),\end{aligned} \tag{51}$$

where V,σ,ψ are given functions with ψ being symmetric with respect to the permutation of either x,y or x_1,y_1, and the index of stability α is a fixed number in the interval (1, 2) (we do not consider here the indices $\alpha\le 1$, as the technique applied below would require a modification in this case). The first (respectively the second) part of (51) stands for the potential (respectively the jump type) interaction.

Remark. The case $\alpha=2$ corresponds to the standard diffusions and is omitted, though all results below are still valid for $\alpha=2$. In particular, in case $\alpha=2$, $\sigma=1$, $\psi=0$ the corresponding Markov model of interacting particles is given by the system of Ito's equations

$$dX_i=dW_i+\sum_j V(X_i,X_j)dt,\qquad i=1,2,\cdots$$

Equations (25) and (28) for generators (50), (51) take form

$$\begin{aligned}\frac{d}{dt}(g,\mu_t)&=-(\sigma(x)|\Delta|^{\alpha/2}g,\mu_t)+\int(V(x,y),\nabla g(x))\mu_t(dx)\mu_t(dy)\\ &+\int(g(x+z)-g(x))\psi(x,y;z)dz\mu_t(dx)\mu_t(dy)\end{aligned} \tag{52}$$

and

$$\frac{d}{dt}f_t(x) = -|\Delta|^{\alpha/2}(\sigma(x)f_t(x)) - \int[(V(x,y),\nabla)f_t(x)f_t(y) + div_x V(x,y)f_t(x)f_t(y)]dy,$$
$$+\int(f_t(x-z)\psi(x-z,y;z) - f_t(x)\psi(x,y;z))f_t(y)dydz \qquad (53)$$

respectively, where

$$div_x V(x,y) = \sum_{i=1}^d \frac{\partial V_i}{\partial x_i}(x,y), \qquad \psi(x,y;z) = \int \tilde{\psi}(x,y;z,w)dw,$$

V_i and x_i being, of course, the coordinates of the vectors V and x. It is possible to get a well posedness of (53) as a consequence of Theorem 3.1. However, we assumed here that $\sigma(x)$ is not affected by interaction, which would lead to better results (arbitrary initial conditions), than those obtained through Theorem 3.1. As a preliminary step we need the basic properties of the corresponding linear problem given by the following theorem that is of independent interest.

Theorem 4.1 *Suppose $\alpha \in (1,2)$, $\sigma(x)$, $A_t(x)$ (respectively $B_t(x)$, $\phi_t(x)$) are continuously differentiable (respectively continuous) functions of x with σ, ϕ, being non-negative and A_t, B_t, ϕ_t depending continuously (in fact, measurably is enough here) on $t \geq 0$, and such that*

$$\sup_{t\leq t_0} \max(\sigma(x), \sigma^{-1}(x), |\nabla\sigma(x)|, |A_t(x)|, |\nabla A_t(x)|, |B_t(x)|, \phi_t(x,z)(1+|z|^{\beta+d})) \leq C_0(t_0) \qquad (54)$$

for some constants $\beta \in (0,2)$, $C_0(t_0) > 0$ and all x, z. Then

(i) the equation

$$\frac{d}{dt}G_t(x) = -\sigma(x)|\Delta|^{\alpha/2}G_t(x) - (A_t(x), \nabla G_t(x)) + B_t(x)G_t(x)$$
$$+\int(G_t(x+z) - G_t(x))\phi_t(x,z)\,dz \qquad (55)$$

has a Green function $G(t,s,x,y)$, $t_0 \geq t \geq s \geq 0$, i.e. its solution with the initial condition $G(s,s,x,y) = \delta(x-y)$, such that

$$G(t,s,x,y) = S_\alpha(t,s,x-y)(1 + O(1)\min(1,|x-y|) + O(t^{1/\alpha}))$$
$$+O(t)(1+|x-y|^{d+\min(\alpha,\beta)})^{-1}, \qquad (56)$$

where $O(1), O(t^{1/\alpha})$ depend only on t_0 , $C_0(t_0)$,

$$S_\alpha(t,s,x-y) = (2\pi)^{-d}\int \exp\left\{-t\sigma(y)|p|^\alpha + i(p,x,-y-\int_s^t A_r(y)dr)\right\}dp \qquad (57)$$

is a shifted stable density (with the index α and the uniform spectral measure) and where the last term in (56) can be omitted whenever $\beta \geq \alpha$;

(ii) $G(t,s,x,y)$ *is everywhere non-negative and satisfies the Chapman-Kolmogorov G equation; moreover, in case* $B_t = 0$ *one has* $\int G(t,s,x,y)dy = 1$ *for all* x *and* $t > s$*;*

(iii) $G(t,s,x,y)$ *is continuously differentiable in* t,s,x,y *whenever* $t > s$ *and*

$$\left|\frac{\partial G}{\partial x}(t,s,x,y)\right| = O((t-s)^{-1/\alpha})G(t,s,x,y) \tag{58}$$

uniformly for $0 \le s < t \le t_0$, $x, y \in R_d$ *;*

(iv) for any $G_s \in C_\infty(X)$ *there exists a unique (classical) solution* G_t *in* $C_\infty(X)$ *to the Cauchy problem of equation* (55) *(i.e. a continuous mapping* $t \to G_t \in C_\infty(X)$ *that solves* (55) *for* $t > s$ *and coincides with* G_s *at* $t = s$*); moreover* $G_t \in C^1(X)$ *for all* $t > s$ *with*

$$\|G_t\|_{C^1(X)} = O(t-s)^{-1/\alpha}\|G_s\|_{C(X)};$$

and if , $\sigma, A, B_t, \phi_t \in C^k(X)$, $k > 0$, *the mapping* $G_s \mapsto G_t$ *is a bounded operator in* $C^k(X)$ *uniformly for* $t \in [0, t_0]$.

Proof. In case $B_t(x) = 0$ and $A_t(x) = A(x)$ being time independent this result is obtained in [32] (the existence of the Green function in case $\psi = 0$ being previously obtained in [31]). It remains to observe that the proof from [32] generalizes straightforwardly to the present non-homogeneous situation.

Remark. In [32] one can find also two-sided estimates for the Green function (56).

Corollary 3 *Under the assumptions of Theorem 4.1 the mapping* $G_s \mapsto G_t$ *extends to (i) the bounded linear mapping* $\mathcal{M}(X) \mapsto \mathcal{M}(X)$ *that is also continuous in the* $*$*-weak topology and is such that its image always has a density (with respect to Lebesgue measure) that solves equation* (55) *for* $t > s$*; and to (ii) the bounded linear mapping* $(C^1_\infty(X))^* \mapsto (C^1_\infty(X))^*$ *that is also continuous in the* $*$*-weak topology.*

Proof. Follows from Theorem 4.1 (iv) and the duality arguments.

The basic properties of the solution to the Cauchy problem of equations (52), (53) are collected in the following theorem.

Theorem 4.2 *Suppose* $\alpha \in (1,2)$, $\sigma(x)$ *is a positive function such that* σ^{-1} *is bounded,* $\sigma \in C^1(X)$, $V \in C^1(X)$ *as a function of the first variable with bounds being uniform with respect to the second one, and a continuous function* ψ *enjoys the bound*

$$\psi(x,y;z) \le \frac{C}{1+|z|^{\beta+d}} \tag{59}$$

for all x, z, y *with some constants* $\beta \in (0,2)$, $C > 0$. *Then the following holds:*

(i) For arbitrary non-negative $f_0 \in C_\infty \cap L_1(X)$ *there exists a unique non-negative classical solution* $f_t \in C_\infty \cap L_1(X) \cap C^1(X)$ *of the Cauchy problem to* (53) *(i.e.* $t \mapsto f_t$ *is a continuous function* $\mathbf{R}_+ \mapsto C_\infty(X)$ *that satisfies* (53) *for* $t > 0$*).*

(ii) The mapping $f_0 \mapsto f_t$ *extends to a strongly continuous semigroup of (nonlinear) continuous isometries of* $L_1^+(X)$ *(i.e.* $(t, f_0) \mapsto f_t$ *is a continuous mapping of two variables with* f_0 *and* f_t *considered in the Banach topology of* $L_1(X)$*).*

(iii) The mapping$f_0 \mapsto f_t$ extends to a semigroup of Lipshitz continuous (in the norm topology) isometries $\mu_0 \mapsto \mu_t$ in $\mathcal{M}^+(X)$ such that (52) holds for any $g \in C^2(X) \cap C_\infty(X)$ and the measure μ_t has a density $f_t \in L_1^+(X) \cap C_\infty(X)$ for all $t > 0$.

(iv) The mapping $\mu_0 \mapsto \mu_t$ is infinitely differentiable with respect to μ_0 and for each k the (signed) measure $\frac{\delta^k \mu_t}{\delta \mu_0^k}(x_1, \dots, x_k; \mu_0)$ is well defined, is uniformly bounded for μ_0 from any bounded set and is continuously differentiable in $x_1, \dots, x_k$. Moreover, if $\sigma(x)$, $V(x,y)$, $\psi(x,y;z)(1+|z|^{\beta+d})$belong to $C^2(X)$ as a function of x uniformly with respect to other variables, then each measure $\frac{\delta^k \mu_t}{\delta \mu_0^k}(x_1, \dots, x_k; \mu_0)$ depends two times continuously differentiable on $x_1, \dots, x_k$ in the topology of the Banach space $(C^2(X))^*$. In particular,

$$\left\| \nabla^2 \left(g, \frac{\delta^2 \mu_t}{\delta \mu_0^2} \right) \right\| \le \|g\|_{C^2(X)} C(t, \|\mu_0\|)$$

with some constant $C(t, \|\mu_0\|)$ for all two times continuously differentiable functions g.

(v) Under the assumptions of (iv) the mapping $f_0 \mapsto f_t$ extends to a semigroup of Lipshitz continuous (in the norm topology) mappings $\mu_0 \mapsto \mu_t$ in $(C^1(X))^*$ such that

$$\|\mu_t\|_{\mathcal{M}(X)} \le C(t) t^{-1/\alpha} \|\mu_0\|_{(C^1(X))^*}.$$

Remark. As is easily seen, the above mapping $t \mapsto \mu_t$ is not continuous in $\mathcal{M}^+(X)$ (unlike its restriction to $L_1(X)$) in the norm topology as a function of t at $t = 0$ for arbitrary μ_0 which is not absolutely continuous with respect to Lebesgue measure. Hence the necessity to work in the weak topology of $M(X)$, which will be exploited in the next Theorem.

Proof. Our strategy in dealing with equation (53) is the same as in the previous section and is based on the observation that a solution to this equation is a fixed point of a mapping $u_t \mapsto F_t$, where for $u_t \in C([0,t_0], \mathcal{M}^+(X))$ the function F_t solves the equation

$$\frac{d}{dt} F_t(x) = -|\Delta|^{\alpha/2}(\sigma(x) F_t(x)) - (V_t(x,[u]), \nabla F_t(x)) - div V_t(x,[u]) F_t(x)$$
$$+ \int (F_t(x-z)\psi_t(x-z,z,[u]) - F_t(x)\psi_t(x,z,[u]))) \, dz \qquad (60)$$

with the initial data $F_0 = u_0$, where

$$V_t(x,[u]) = \int V(x,y) u_t(dy), \qquad \psi_t(x,z,[u]) = \int \psi(x,y;z) u_t(dy).$$

In order to solve this equation using Theorem 4.1 we change the unknown function F_t to $G_t(x) = \sigma(x) F_t(x)$ leading to the equation

$$\frac{d}{dt} G_t(x) = -\sigma(x)|\Delta|^{\alpha/2} G_t(x) - (V_t(x,[u]), \nabla G_t(x))$$
$$+ [(V_t(x,[u]), \nabla \ln \sigma(x)) - div V_t(x,[u])] G_t(x)$$
$$+ \int (\sigma(x)\sigma^{-1}(x-z) G_t(x-z) \psi_t(x-z,z,[u]) - G_t(x)\psi_t(x,z,[u])) dz dy. \qquad (61)$$

Applying Theorem 4.1 to equation (61) yields a solution $F_t = U(t,s)F_s$ to the Cauchy problem of equation (60) in the form

$$F_t(x) = \int \sigma^{-1}(x) G(t,s,x,y;[u]) \sigma(y) F_s(dy),$$

which is unique for F_s being from $C_\infty(X)$ and where $G(t,s,x,y;[u])$ is the Green function of equation (61). As is also the case in Theorem 3.1 equation (60) is dual to a similar equation, but which unlike (60) preserves constants. Hence the family of the operators $U(t,s)$ extends to a strongly continuous semigroups of isometries in $L_1^+(X)$ and $\mathcal{M}^+(X)$, where strong continuity in case $\mathcal{M}(X)$ is understood in the sense of the weak topology of $\mathcal{M}(X)$. The key property of this semigroup follows from (58): $F_t \in W^1(X)$ and

$$\|F_t\|_{W^1(X)} \le C_1(t_0)(t-s)^{-1/\alpha} \|F_s\|_{\mathcal{M}(X)} \tag{62}$$

whenever $t > s$ and $F_s \in \mathcal{M}(X)$.

We shall follow now the same line of argument as in Theorem 3.1 but with u_0 being an arbitrary element of $L_1^+(X)$ or $\mathcal{M}^+(X)$. Instead of (41) one has

$$\|(L_1(s) - L_2(s))g\|_{L_1(X)} \le C_1 \|u^1(s) - u^2(s)\|_{\mathcal{M}(X)} \|g\|_{W_1(X)}. \tag{63}$$

Consequently, (40), (62), (63) yield

$$\sup_{0<s\le t} \|F_s([u^1]) - F_s([u^2])\|_{L_1(X)} \le t^{1-1/\alpha} C_1 C_2(t_0) \|u_0\|_{\mathcal{M}(X)} \sup_{0<s\le t} \|u^1(s) - u^2(s)\|_{L_1(X)}. \tag{64}$$

Hence the mapping u_t to F_t in the subset of $S_0C([0,t_0], \mathcal{M}^+(X))$ obtained by fixing u_0 is a contraction whenever

$$t^{1-1/\alpha} C_1 C_2(t_0) \|u_0\|_{\mathcal{M}(X)} < 1. \tag{65}$$

Consequently one obtains a solution to (53) on the interval $[0,t_1]$ for any $t_1 \le t_0$ satisfying (65), which is unique for $u_0 \in C_\infty(X)$. If $t_1 < t_0$ one can repeat this procedure starting from the time t_1 obtaining a unique continuation of the solution to the interval $[0,t_2]$ with $t_2 - t_1$ satisfying (65), etc.

Let us show the required continuity properties of the constructed solution. Suppose $u^1(t), u^2(t)$ are the fixed points to $u_t \mapsto F_t$ (i.e. the constructed solutions to equation (53)) with different initial values $u_0^1, u_0^2 \in \mathcal{M}(X)$ both satisfying (65). Let $U^i(t,s), i = 1,2$, denote the corresponding propagators to (60). Then

$$\begin{aligned} u^1(t) - u^2(t) &= U^1(t,0)u_0^1 - U^2(t,0)u_0^2 = (U^1(t,0) - U^2(t,0))u_0^1 + U^2(t,0)(u_0^1 - u_0^2) \\ &= \int U^2(t,s)(L_s^1 - L_s^2)U^1(s,0)u_0^1 \, ds + U^2(t,0)(u_0^1 - u_0^2). \end{aligned} \tag{66}$$

Consequently

$$\sup_{0\le s\le t} \|u^1(s) - u^2(s)\|_{\mathcal{M}(X)} \le \omega \sup_{0\le s\le t} \|u^1(s) - u^2(s)\|_{\mathcal{M}(X)} + \|u_0^1 - u_0^2\|_{\mathcal{M}(X)}$$

with a $\omega \in (0,1)$ (being given by the l.h.s. of (65)), and thus

$$(1-\omega) \sup_{0\le s\le t} \|u^1(s) - u^2(s)\|_{\mathcal{M}(X)} \le \|u_0^1 - u_0^2\|_{\mathcal{M}(X)}.$$

This implies the continuity (even Lipshitz continuity) of the mapping $\mu_0 \mapsto \mu_t$ in the norm topology both in $L_1^+(X)$ and $\mathcal{M}(X)$.

The continuity of $f_0 \mapsto f_t$ as a $L_1(X)$-valued function of t follows from the same property of the linear problem.

Statement (iv) follows from the possibility to differentiate the equation for μ_t with respect to the initial condition arbitrary number of times, as the r.h.s. of the equation depends on μ_t quadratically. Differentiation leads to the time non-homogeneous equation of type (55), whose well-posedness follows again from Theorem 4.1.

To get (v) one extends the above fixed point arguments from $\mathcal{M}(X)$ to $(C^1(X))^*$ using instead of (63) and (64) the inequalities

$$\|(L_1(s) - L_2(s))g\|_{(C^1(X))^*} \le C_1 \|u^1(s) - u^2(s)\|_{(C^1(X))^*} \|g\|_{L_1(X)}$$

and respectively

$$\sup_{0<s\le t} \|F_s([u^1]) - F_s([u^2])\|_{(C^1(X))^*}$$
$$\le t^{1-1/\alpha} C_1 C_2(t_0) \|u_0\|_{(C^1(X))^*} \sup_{0<s\le t} \|u^1(s) - u^2(s)\|_{(C^1(X))^*}.$$

We shall apply the above results to the study of the semigroup of positivity preserving contractions $T_t F(\mu) = F(\mu_t)$ on the space of bounded functions on $\mathcal{M}_M^+(X)$, which is a subset in $\mathcal{M}^+(X)$ consisting of measures with norms bounded by M. Let $C_0(\mathcal{M}_M^+(X))$ denote the closure in $C(\mathcal{M}^+(X))$ (the space of measures is considered in the $*$-weak topology) of the set of polynomial functionals $C_{\infty,fin}^{sym}(\mathcal{X})$ consisting of finite linear combinations f of the functionals of the form

$$F(\mu) = \frac{1}{m!} \int g^m(x_1, \ldots, xm)\mu(dx_1) \cdots \mu(dx_m) = (g^m, \mu^{\tilde{\otimes} m}) \tag{67}$$

with $g^m \in C_\infty^{sym}(X^m)$. As polynomial functionals are known to be dense in $C(\mathcal{M}_M^+(Z))$ for compact spaces Z (which is a direct consequence of the Weierstrass theorem), it follows that for a locally compact metric space X the space $C_0(\mathcal{M}_M^+(X))$ consists of elements F from $C(\mathcal{M}_M^+(X))$ such that $F_n \to F$ in $C(\mathcal{M}_M^+(X))$, where $F_n(\mu) = F(\chi_n \mu)$ (χ_n being the indicator function of the ball of radius n with some fixed center).

Recall that the variational derivative $\frac{\delta F}{\delta \mu}$ (or shortly δF) of a functional $F \in C(\mathcal{M}(X))$ is defined as

$$\delta F(x, \mu) = \lim_{h \to 0, h>0} \frac{1}{h}(F(\mu + h\delta_x) - F(\mu)).$$

The space $C^1(\mathcal{M}_M^+(X))$ is defined as the spaces of functionals such that δF exists for all $x \in X$ and $\mu \in \mathcal{M}_M^+(X)$ and is a continuous function of two variables there. One sees by inspection that for $F \in C^1(\mathcal{M}_M^+(X))$ and arbitrary $\mu, \nu \in \mathcal{M}_M^+(X)$ the formula

$$F(\nu) - F(\mu) = \int_0^1 ds \int \delta F(x, \mu + s(\nu - \mu))(\nu - \mu)(dx)$$

holds and that the space $C_0(\mathcal{M}_M^+(X))$ contains the space $C_0^1(\mathcal{M}^+(X))$ of functionals from $C^1(\mathcal{M}_M^+(X))$ such that $\delta F(\cdot, \mu) \in C_\infty(X)$ uniformly for $\mu \in \mathcal{M}_M^+(X)$.

Theorem 4.3 *Under the assumptions of Theorem 4.2 suppose additionally that $V(x,y)$ and $\psi(x,y;z)(1+|z|^{d+\beta})$ belong to $C_\infty(X) \cap C^1(X)$ as functions of y uniformly with respect to other variables. Then* *(i) the mapping $(\mu_0, t) \mapsto \mu_t$ is a continuous mapping of two variables with μ_0 , μ_t considered in the weak topology;*

(ii) the family of linear operators $T_t F(\mu) = F(\mu_t)$ defines a contraction semigroup on $C(\mathcal{M}^+(X))$ and on $C(\mathcal{M}^+_M(X))$ for arbitrary M , which is strongly continuous on $C_0(\mathcal{M}^+_M(X))$;

(iii) the subspace $C^{1,2}_0(\mathcal{M}^+_M(X))$ of $C^1_0(\mathcal{M}^+_M(X))$ consisting of functionals F with $\delta F(x,\mu)$ being two times continuously differentiable in x with the first and second derivatives belonging to $C_\infty(X)$ uniformly for $\mu \in \mathcal{M}^+_M(X)$ represents an invariant core of the generator $\mathcal{L}_B$ of the semigroup T_t on $C_0(\mathcal{M}^+_M(X))$;

(iv) the space of polynomial functionals $C^{2,sym}_{\infty,fin}(X)$ consisting of linear combinations of the functionals of form (67) *with $g \in C_\infty(X)$ being two times differentiable with its first and second derivatives belonging to $C_\infty(X)$ is also a core of the generator $\mathcal{L}_B$.*

Proof. (i) The continuity of the mapping $t \to \mu_t$ in the weak topology follows from duality. Let us show that if $u^n_0 \in \mathcal{M}^+(X)$ converges weakly to $u_0 \in \mathcal{M}^+(X)$, $n \to \infty$, then $u^n(t)$ converges to $u(t)$ weakly for each $t > 0$, where $u^n(t)$, $n = 1, 2, \dots, u(t)$ are the fixed points of the mapping $u_t \mapsto F_t$ (considered in the proof of the previous theorem) with initial data u^n_0, u_0 respectively. Let $U^n(t,s)$ and $U(t,s)$ denote the propagators corresponding to $u^n(t)$ and $u(t)$. Notice that the adjoint operator to the operator L_t from the r.h.s. $L_t(F_t)$ of (60) equals

$$L^*_t g = -\sigma(x)|\Delta|^{\alpha/2} g + (V_t(x,[u]), \nabla F_(t)) + \int (g(x-z) - g(x))\psi_t(x,z,[u]))dz. \qquad (68)$$

Denoting

$$\gamma^n_t(x) = V_t(x,[u^n - u]) = \int V_t(x,y)(u^n(t) - u(t))(dy),$$
$$\delta^n_t(x,z) = \psi(x,z,[u^n - u])(1 + |z|^{d+\beta})$$

(we use the notations introduced in the proof of the previous theorem) and using (66) yields

$$\gamma^n_t(x) = \left(\int_0^t (U^n_{s,0})^*(L^*_{n,s} - L^*_s)U^*_{t,s}V(x,\cdot)\,ds, u^n_0\right) + (U^*_t(t,0)V(x,\cdot), u^n_0 - u_0),$$
$$\delta^n_t(x,z) = \left(\int_0^t (U^n_{s,0})^*(L^*_{n,s} - L^*_s)U^*_{t,s}\psi(x,\cdot;z)(1+|z|^{d+\beta})ds, u^n_0\right)$$
$$+ (U^*_t(t,0)\psi(x,\cdot;z)(1+|z|^{d+\beta}), u^n_0 - u_0).$$

Hence

$$|\gamma^n_t(x)| \le K\left(\int_0^t (\sup_x |\gamma^n_s(x)| + \sup_{x,z} |\delta^n_s(x,z)|)ds, u^n_0\right) + |(U^*_t(t,0)V(x,\cdot), u^n_0 - u_0)|$$

for some constant K with a similar estimate for δ_t yielding for

$$\xi_t = \sup_x |\gamma^n_t(x)| + \sup_{x,z} |\delta^n_t(x,z)|$$

the estimate

$$\xi_t \le K \int_0^t \xi_s ds + \sup_x |(U_t^*(t,0)V(x,\cdot), u_0^n - u_0)| + \sup |(U_t^*(t,0)\psi(x,\cdot,z)(1+|z|^{d+\beta}), u_0^n - u_0)|$$

(with some other constant K). As the sets of functions $V(x,\cdot)$ and $\psi(x,\cdot;z)(1+|z|^{d+\beta})$ are compact in C_∞ (bounded with uniformly bounded derivatives), the last two terms on the r.h.s. of this inequality tend to zero as $n \to \infty$ uniformly for $t \in [0,T]$ for arbitrary $T > 0$. Consequently (due to Gronwall's lemma) the same holds for ξ_t, and hence for γ_t and δ_t. Consequently, (66) implies

$$(g, u^n(t) - u(t)) = O(1) \sup_{s\in[0,t]} \xi_s \int_0^t s^{-1/\alpha} ds + |(U_t^*(t,0)g, u_0^n - u_0)|$$

for arbitrary $g \in C_\infty$, which implies the convergence $(g, u^n(t) - u(t)) \to 0$ and hence the weak convergence $u^n(t) \to u(t)$.

(ii) By (i) the semigroup T_t preserves the spaces $C(\mathcal{M}_M^+(X))$ and $C(\mathcal{M}^+(X))$. Let us show that the space $C_0(\mathcal{M}_M^+(X))$ is invariant under the action of this semigroup. To do this one needs to show that the functional $F(\mu_t)$ belongs to $C_0(\mathcal{M}_M^+(X))$ for any $t > 0$ and F of form (67). For simplicity let us give the arguments for the case $k = 1$ only, denoting g^1 by g (the general case is quite similar). Let χ_n denote a smoothed indicator function of the ball of radius n in X (a continuous function that equals one in this ball, vanishes outside the ball of radius $n+1$ and is bounded by one everywhere). Since for a compact X the polynomial functionals are known to be dense in $C(\mathcal{M}_M^+(X))$ (which M is an easy consequence of the Weierstrass theorem), the functional $(g, \mu_t(\chi_n \mu_0))$ belongs to $C_0(\mathcal{M}_M^+(X))$ for any n. Hence it remains to show that $(g, \mu_t(\chi_n \mu_0))$ converges to $(g, \mu_t(\mu_0))$ for arbitrary t uniformly for all $\mu_0 \in \mathcal{M}_M(X)$. Again using (66) one writes

$$(g, \mu_t(\chi_n\mu_0) - \mu_t(\mu_0)) = \left(\int_0^t (U_{s,0}^n)^*(L_{n,s}^* - L_s^*)U_{t,s}^* V(x,\cdot)ds, u_0^n \right) + (g, U(t,0)(\chi_n\mu_0 - \mu_0)). \tag{69}$$

The second term here is uniformly small, because $g \in C_\infty(X)$, and $U_{t,0}^*$ preserves this property uniformly, as the coefficients of all generators are uniformly bounded due to the assumption that $|\mu_0| \le M$. For the first term this is not directly obvious. But we can use the trick from the part (i). Namely, consider first the functions g of form $V(x,\cdot)$ and $\psi(x,\cdot;z)(1+|z|^{d+\beta})$ and prove the smallness using Gronwall's lemma. When this is done the smallness of the first term in (69) becomes clear.

It remains to show that T_t is strongly continuous on polynomial functionals of the form (67). Again let us reduce the discussion to the case $k = 1$ only. Thus we need to show that (g, μ_t) tends to (g, μ_0) as $t \to 0$ for any given $g \in C_\infty(X)$ uniformly for μ_0 with a bounded norm. For this it is enough to prove that

$$\sup_{\|\mu_0\| \le M} \|U(t,0)^* g - g\|_{C_\infty(X)} \to 0,$$

where $U(t,0)$ is the propagator corresponding to μ_0. But this follows again from the observation that the coefficients of the corresponding generators are uniformly bounded for $\|\mu_0\| \le M$.

(iii) Equation (52) yields

$$\frac{d}{dt}F(\mu_t = -\int \sigma(x)|\Delta|^{\alpha/2}\delta F(x,\mu_t)\mu_t(dx) + \int\int (V(x,y), \nabla_x \delta F(x,\mu_t))\mu_t(dx)\mu_t(dy)$$
$$+\int (\delta F(x+z,\mu_t) - \delta F(x,\mu_t))\psi(x,y;z)\, dz\mu_t(dx)\mu_t(dy) \tag{70}$$

for $F \in C_0^{1,2}(\mathcal{M}_M^+(X))$. To prove that F belongs to the domain of $\mathcal{L}_B$, one needs to show that$(F(\mu_t) - F(\mu_0))/t$ has a limit as $t \to 0$ uniformly for all μ_0 from $\mathcal{M}_M^+$. The representation

$$F(\mu_t) - F(\mu_0) = t\frac{d}{dt}\mid_{s=0} F(\mu_s) + \int_0^t \left(\frac{d}{ds}F(\mu_s) - \frac{d}{ds}\mid_{s=0} F(\mu_s)\right) ds,$$

implies that it is sufficient to show that $\frac{d}{dt}F(\mu_t) - \frac{d}{ds}\mid_{s=0} F(\mu_s)$ tends to zero as $t \to 0$ uniformly for μ_0 from $\mathcal{M}_M^+$. But due to (70), this difference can be written as a linear combination (with bounded coefficients) of the differences of the form $(\phi, \mu_t - \mu_0)$ with given functions ϕ from C_∞ and the integrals $\int(\delta F(x,\mu_t) - \delta F(x,\mu_0))\mu_0(dx)$. The first differences tend to zero by the part (ii) of the Theorem. The integral is small because of the continuity of $\delta F(x,\mu)$ in μ (uniform on x and μ from compact sets). By Theorem 4.2 (iv) the space $C_0^{1,2}(\mathcal{M}_M^+(X))$ is invariant under the action of the semigroup T_t and hence represents a core.

(iv) Follows from the possibility to approximate the functionals from $C_0^{1,2}(\mathcal{M}_M^+(X))$ (together with their derivatives) by polynomials, which follows from Proposition A.1 of the Appendix.

5 The law of large numbers

Let $h > 0$ be a positive parameter. We shall denote by δ_x the Dirac measure at x. A key role in the theory of measure-valued limits is played by the inclusion

$$x = (x_1, \dots, x_l) \mapsto h\delta_x = h(\delta_{x_1} + \dots + \delta_{x_l}), \tag{71}$$

which defines a homeomorphism between $S\mathcal{X}$ and the set $\mathcal{M}_{\delta,h}^+(X)$ of h-scaled finite sums of δ-measures.

Let X and D be the same as in Theorem 2.1, and let $B_1, \dots, B_K$ be a collection of conditionally positive operators in $C_\infty(X), \dots, C_\infty(X^K)$ respectively that generate Feller semigroups and, in particular, have representation (20). A *process $Z_B(t)$ of K-ary interaction in $\mathcal{X}$*, where any collection of $k \le K$ particles interact according to B_k, is defined through its generator

$$L_B f(x_1, \dots, x_l) = \sum_{k=1}^K L_k f(x_1, \dots, x_l) = \sum_{k=1}^{\min(K,n)} \sum_{I \subset \{1,\dots,l\}:|I|=k} (B_k^I f)(x_1, \dots, x_l) \tag{72}$$

(see notations for B_k^I in Theorem 2.5).

We shall study now a scaling limit of the large number of particles where k-ary part L_k of the generator is scaled by the factor h^{k-1} and $\mathbf{x} = (x_1, \dots, x_l)$ is substituted by a measure according to (71). Denoting $F(h\delta_x) = f(\mathbf{x})$ for $f \in C^{sym}(\mathcal{X})$ yields the scaled generator on$C(\mathcal{M}^+_{\delta,h}(X))$ given by

$$(L^h_B F)(h\delta_{\mathbf{x}}) = \sum_{k=1}^{K} h^{k-1} L^h_k F(h\delta_{\mathbf{x}}) \tag{73}$$

with

$$L^h_k F(h\delta_{\mathbf{x}}) = (L_k f)(\mathbf{x}). \tag{74}$$

For a linear operator B_k in $C^{sym}(X^k)$ let $\tilde{B}^l_k, l = 1, 2, \dots,$ denote the linear operators $C^{sym}(X^l) \to C^{sym}(X^k)$ defined as

$$(\tilde{B}^l_k g(\mathbf{x}) = \sum_{q=1}^{l} (-1)^{l-q} \int_{X^{l-q}} B^{\mathbf{x}}_k \left(\int_{X^q} g(\mathbf{y}, \mathbf{z}) \delta^{\tilde{\otimes} q}_{\mathbf{x}}(d\mathbf{y}) \right) \delta^{\tilde{\otimes}(l-q)}_{\mathbf{x}}(d\mathbf{z}),$$

where of course $\mathbf{x} \in X^k$, $\mathbf{y} \in X^q$, $\mathbf{z} \in X^{l-q}$ and $B^{\mathbf{x}}_k$ means the action of B_k on the variable $\mathbf{x}$. In particular,

$$(\tilde{B}^1_k g(\mathbf{x}) = B_k \left(\int g(y) \delta_{\mathbf{x}}(dy) \right) = (B_k g^+)(\mathbf{x}), \tag{75}$$

$$(\tilde{B}^l_k g(\mathbf{x}) = \sum_{q=1}^{l} \frac{(-1)^{l-q}}{q!(l-q)!} (B^x_1 g(x, \dots, x, z, \dots, z)) \mid_{z=x}, \tag{76}$$

where x is written q times and z is written $l - q$ times in the last formula.

Proposition 5.1 *Let all B_k be conservative, i.e. $B_k 1 = 0$. Let $F(\mu)$ be given by* (67) *with $g = g^m \in C^{2,sym}_\infty(X^m)$ and let $\mathbf{x} = (x_1, \dots, x_n)$ with $n \geq \max(2, k)$ be chosen. Then*

$$h^{k-1} L^h_k F(h\delta_{\mathbf{x}}) = \sum_{l=1}^{m} h^{l-1} \int_{X^{m-l}} \left(\Phi^k_h [\tilde{B}^l_k g_{\mathbf{w}}](h\delta_{\mathbf{x}}) \right) (h\delta_{\mathbf{x}})^{\tilde{\otimes}(m-l)}(d\mathbf{w}), \tag{77}$$

where $g_{\mathbf{w}}(z) = g(\mathbf{w}, \mathbf{z})$ and $\Phi^k_h[\tilde{B}^l_k g_{\mathbf{w}}]$ is given by Proposition B.1 of Appendix.

Proof. By (72)-(74)

$$h^{k-1} L^h_k F(h\delta_{\mathbf{x}}) = \frac{1}{m!} h^{m+k-1} \sum_{I \subset \{1,\dots,n\}: |I|=k} B^I_k \sum_{j_1,\dots,j_m=1} g(x_{j_1}, \dots, x_{j_m}),$$

where B^I_k means the action of B_k on the variables x_I. Denoting by q the number of indexes in $\{j_1, \dots, j_m\}$ that belong to I one rewrites it as

$$\frac{h^{m+k-1}}{m!} \sum_{I \subset \{1,\dots,n\}: |I|=k} B^I_k \sum_{q=1}^{m} C^q_m \sum_{j_1,\dots,j_q \in I} \sum_{i_1,\dots i_{m-q} \notin I} g(x_{j_1}, \dots, x_{j_q}, x_{i_1}, \dots x_{i_{m-q}})$$

$$= \frac{h^{m+k-1}}{m!} \sum_{q=1}^{m} C^q_m \sum_{I \subset \{1,\dots,n\}: |I|=k} B^I_k \int_{X^q} \int_{X^{m-q}} g(\mathbf{y}, \mathbf{z}) \delta^{\otimes q}_{\mathbf{x}_I}(d\mathbf{y}) (\delta_{\mathbf{x}} - \delta_{\mathbf{x}_I})^{\otimes(m-q)}(d\mathbf{z})$$

$$= \frac{h^{m+k-1}}{m!} \sum_{q=1}^{m} C^q_m \sum_{p=0}^{m-q} C^p_{m-q} (-1)^p \sum_{I \subset \{1,\dots,n\}: |I|=k}$$

$$\int \int \left(B^I_k \int g(\mathbf{y}, \mathbf{z}, \mathbf{w}) \delta^{\otimes q}_{\mathbf{x}_I}(d\mathbf{y}) \right) \delta^{\otimes p}_{\mathbf{x}_I}(d\mathbf{z}) \delta^{\otimes(m-q-p)}_{\mathbf{x}}(d\mathbf{w}),$$

which by changing the index p to the index $l = q + p$ and using $C_m^q C_{m-q}^p = C_m^l C_l^p$ reads as

$$\frac{h^k}{m!} \sum_{l=1}^m h^{l-1} C_m^l \sum_{q=1}^l C_l^q (-1)^{l-q} \sum_{I \subset \{1,\dots,n\}:|I|=k}$$
$$\int \left[\int B_k^I \left(\int g(\mathbf{y},\mathbf{z},\mathbf{w}) \delta_{\mathbf{x}_I}^{\otimes q}(d\mathbf{y}) \right) \delta_{\mathbf{x}_I}^{\otimes (l-q)}(d\mathbf{z}) \right] (h\delta_{\mathbf{x}})^{\otimes (m-l)}(d\mathbf{w}),$$

which yields (77) by (114).

As for F of form (67) one has

$$\frac{\delta^l F}{\delta Y^l}(x_1,\dots,x_l;Y) = (g^m(x_1,\dots,x_l,\cdot), Y^{\tilde{\otimes}(m-l)})$$

it follows from (77) that one can extend the operator L_k^h to act on more general infinitely differentiable functions $F(Y)$ of measures by the formula

$$h^{k-1} L_k^h F(Y) = \sum_{l=1}^\infty h^{l-1} \Phi_h^k \left[\tilde{B}_k^l \frac{\delta^l F}{\delta Y^l}(Y) \right](Y). \tag{78}$$

As a direct consequence of (78) one obtains also the following

Proposition 5.2 *Let all B_k be conservative, i.e. $B_k 1 = 0$. Let $F(Y) = (g, Y^{\tilde{\otimes}})$ with $g \in C_{\infty,fin}^{sym}(\mathcal{X}) \cap C^2(\mathcal{X})$ and thus*

$$f(\mathbf{x}) = F(h\delta_{\mathbf{x}}) = (g, (h\delta_{\mathbf{x}})^{\tilde{\otimes}}) = \sum_{m=1}^M \frac{h^m}{m!} \sum_{i_1,\dots,i_m=1}^l g(x_{i_1},\dots,x_{i_m}) \tag{79}$$

for $\mathbf{x} = (x_1,\dots,x_l)$. Then

$$(L_B^h F)(h\delta_{\mathbf{x}}) = (\mathcal{L}_B g, (h\delta_{\mathbf{x}})^{\tilde{\otimes}}) + O(h), \tag{80}$$

where $\mathcal{L}_B$ is given by (35) and $O(h)$ depends on M (the maximal non-vanishing component of g) and $\|g\|_{C^2(\mathcal{X})}$.

The analytic approach to proving the law of large numbers is described by the following result.

Theorem 5.1 *Suppose the operator (72) (with domain containing $C_c^2(\mathcal{X})$) generates a Feller semigroup T_t^h on $C_\infty^{sym}(\mathcal{X})$, the kinetic equation (10) is well posed and the corresponding evolution defines a strongly continuous semigroup T_t on $C_0(\mathcal{M}_M^+(X))$ with the space $C_{\infty,fin}^{2,sym}(\mathcal{X})$ representing a core for its generator. Then the family of Feller semigroups T_t^h converges strongly to the semigroup T_t as $h \to \infty$. In particular, the family of processes $Z_B^h(t)$ on $\mathcal{M}_{\delta,h}^+(X)$ specified by generator (73) converges weakly to the deterministic evolution described by the kinetic equation whenever the initial conditions converge.*

Proof. Follows from (80) and a well known general result (see e.g. [21] or [47] for different proofs) on the convergence of contraction semigroups. More precisely, the above result states the convergence for the initial conditions of the kinetic equations being equal to that of $Z_B^h(t)$, and our claim then follows from the weak continuity of T_t .

In Section 4 we did all the necessary job to apply this theorem to the system of interacting stable jump diffusions specified by the generators (50), (51). In this case operator (72) becomes

$$Lf(x_1,\dots,x_l) = -\sum_{j=1}^{l}\sigma(x_j)|\Delta_{x_j}|^{\alpha/2}f(x_1,\dots,x_l) + 2\sum_{i<j}\left[V(x_i,x_j)\frac{\partial f}{\partial x_i} + V(x_j,x_i)\frac{\partial f}{\partial x_j}\right.$$
$$\left. + \int (f(x_i+x_1,x_j+x_2) - f(x_i,x_j))\tilde{\psi}(x_i,x_j;x_1,x_2)dx_1dx_2\right], \tag{81}$$

where f in the last integral depends of course on other arguments that are not written explicitly to shorten the formula. Clearly $C_\infty(X^l)$ is invariant under L for each l, and the restriction of L to each $C_\infty(X^l)$ generates a Feller semigroup there (by Theorem 4.1) whenever

$$\tilde{\psi}(x,y;z,w) \le \frac{C}{1+(|z|+|w|)^{2d+\beta}} \tag{82}$$

with some positive C and β (strictly speaking the condition on the density of the integral operator in Theorem 4.1 differs from (82), but the only relevant property of this density is that it defines a bounded integral operator, which is clearly the case under (82). Again by Theorem 4.1 the core of the generator $\mathcal{L}$ of T_t belongs to the domain of L (taken in its representation in $\mathcal{M}^+_{\delta,h}(X)$). This implies the following result.

Theorem 5.2 *Under* (82) *and the assumptions of Theorem 4.3 the operators* (50), (51) *describe a system satisfying all the conditions and hence all the statements of Theorem 5.1.*

As a first step in the direction of the central limit, we shall propose now a method for estimating the rate of convergence in Theorem 5.1, illustrating it on our main model of interacting stable processes. To this end, one needs to estimate series (78). This is not difficult for bounded generators B_k (see examples in Section 7), but becomes impossible in general case. To circumvent this difficulty we shall give an alternative representation for the l.h.s. of (78). For the generators of the jump type, we shall do it for simplicity only for the case $k=1$.

If B_1 is given by the integral part of (1), (2), i.e.

$$(B_1g)(x) = \int (g(x+y) - g(x) - (\nabla g(x), y)\chi(y))\nu(x,dy), \tag{83}$$

then $L_1^hF(h\delta_{\mathbf{x}})$ equals

$$\int\int \left[\frac{1}{h}(F(h\delta_{\mathbf{x}} + h(\delta_{z+y}-\delta_z)) - F(h\delta_{\mathbf{x}})) - \left(\nabla_z \frac{\delta F}{\delta\mu}(z;\mu), y\chi(y)\right)\right](h\delta_{\mathbf{x}})(dz)\nu(z,dy)$$

for $F(h\delta_{\mathbf{x}}) = f(\mathbf{x})$, because clearly

$$\int (f(x_1+y,x_2,\dots,x_n) - f(x_1,\dots,x_n) - (\nabla_{x_1}f(\mathbf{x}), y\chi(y)))\nu(x_1,dy)$$
$$= \int (F(h\delta_{\mathbf{x}} + h(\delta_{x_1+y}-\delta_{x_1})) - F(h\delta_{\mathbf{x}}) - h\left(\nabla_{x_1}\frac{\delta F}{\delta\mu}(x_1;h\delta_{\mathbf{x}}), y\chi(y)\right)\nu(x_1,dy).$$

Hence the operator $L_1^h F$ can be extended by the formula

$$L_1^h F(\mu) = \int\int \left[\frac{1}{h}(F(\mu + h(\delta_{z+y} - \delta_z)) - F(\mu)) - \left(\nabla_z \frac{\delta F}{\delta \mu}(z;\mu), y\chi(y)\right)\right] \mu(dz)\nu(z,dy) \tag{84}$$

to the functions F on measures such that $\frac{\delta F}{\delta \mu}(x;\mu)$ exists, is differentiable in x, and the integral in (84) is well defined. If moreover, $\frac{\delta^2 F}{\delta \mu^2}(x,y;\mu)$ exists and belongs to $C^2(\mathbf{R}^{2d})$ as a function of x,y uniformly for μ from any bounded (in norm) set, then expanding the r.h.s. of (84) in h and using (83) yields

$$L_1^h F(\mu) = \left(B_1 \frac{\delta F}{\delta \mu}, \mu\right)$$
$$+ h \int\int \left(\int_0^1 \frac{\delta^2 F}{\delta \mu^2}(\mu + th(\delta_{z+y} - \delta_z))dt, (\delta_{z+y} - \delta_z)^{\tilde{\otimes} 2}\right) \nu(z,dy)\mu(dz). \tag{85}$$

This is the representation we need to estimate the rate of convergence in Theorem 5.2.

On the other hand, series (78) becomes finite in case of pure differential generators B_k, as in this case the operators $\tilde{B}_k^l$ vanish for large l (which gives an analytic explanation for the fact that interacting diffusions are simpler for analysis). We shall demonstrate this statement only for the most important case of binary interactions.

Proposition 5.3 *(i) If*

$$B_2 g(x,y) = V(x,y)\frac{\partial g}{\partial x}(x,y) + V(y,x)\frac{\partial g}{\partial y}(x,y), \tag{86}$$

then $\tilde{B}_2^l = 0$ for all $l > 1$;

(ii) if

$$B_2 g(x,y) = c(x,y)\frac{\partial^2 g}{\partial x^2} + 2\gamma(x,y)\frac{\partial^2 g}{\partial x \partial y} + c(y,x)\frac{\partial^2 g}{\partial y^2},$$

where c and γ are symmetric matrices and $\gamma(x,y) = \gamma(y,x)$, then $\tilde{B}_2^l = 0$ for all $l > 2$.

Proof. (i) Using the general definition of $\tilde{B}_2^l$ in the present situation yields for $g^l \in C^{sym}(X^l)$

$$\tilde{B}_2^l g^l(x_1,x_2) = \sum_{q=1}^{l} \frac{(-1)^{l-q}}{q!(l-q!} \sum_{j=0}^{q}\sum_{i=0}^{l-q} C_q^j C_{l-q}^i$$
$$\times \left[V(x_1,x_2)\frac{\partial}{\partial x_1} + V(x_2,x_1)\frac{\partial}{\partial x_2}\right] g(x_1,\dots,x_1,x_2,\dots,x_2,z_1,\dots,z_1,z_2,\dots,z_2) \mid_{z_1=x_1,z_2=x_2},$$

where in the last term there are j arguments x_1, $q-j$ arguments x_2, i arguments z_1, and $l-q-i$ arguments z_2. Let us prove that the coefficients at $V(x_1,x_2)$ and $V(x_2,x_1)$ both vanish in case $l > 1$. By symmetry it is enough to deal with the first coefficient only. Hence, we need to prove that

$$\sum_{q=1}^{l} C_l^q (-1)^{l-q} \sum_{j=0}^{q}\sum_{i=0}^{l-q} C_q^j C_{l-q}^i j \frac{\partial g}{\partial x_1}(x,\dots,x,y,\dots,y) = 0$$

($m = i + j$ variables x and $l - m$ variables y) for any differentiable g. Let us show that the coefficient at $\frac{\partial g}{\partial x_1}(x, \dots, x, y, \dots, y)$ (m times x and $l - m$ times y) vanishes for any $m = 1, \dots, l$, i.e. that

$$\sum_{j=1}^{m} \sum_{q=j}^{j+l-m} C_l^q C_q^j C_{l-q}^{m-j} j(-1)^{l-q} = 0.$$

By shifting the index q in each sum, namely denoting $q+m-j$ by q, reduces this equation to

$$\sum_{q=m}^{l} \sum_{j=1}^{m} C_j^{q+j-m} C_{q+j-m}^{j} C_{l-q-j+m}^{m-j} j(-1)^{l-q/m-j} = 0,$$

which after simple algebraic manipulations with binomial coefficients rewrites as

$$\sum_{q=m}^{l} (-1)^{l-q} C_l^q C_q^m \sum_{j=1}^{m} C_m^j j(-1)^{m-j} = 0. \qquad (87)$$

But this holds, as the sum over j vanishes for each fixed $m > 1$, and hence the l.h.s reduces h to $\sum_{q=1}^{l}(-1)^{l-q} q C_l^q$, which again vanishes for $l > 1$.

(ii) The proof is analogous. Omitting the detail note only that instead of (87) one uses here the identity

$$\sum_{q=m}^{l} (-1)^{l-q} C_l^q C_q^m \sum_{j=1}^{m} C_m^j j(j-1)(-1)^{m-j} = 0.$$

Similarly one easily shows that B_1^l vanishes for all $l > 2$ for any second order differential operator B_1, showing the finiteness of series (78) in case of interacting diffusions.

We can prove now the main result of this section.

Theorem 5.3 *Under the assumptions of Theorem 5.2 assume also for simplicity that* $\tilde{\psi} = 0$. *Let* $g^m(x_1, \dots, x_m) \in C^2(X^m)$ *and* $F(\mu) = (g^m, \mu^{\tilde{\otimes} m})$. *Assume* $\mu = h\delta_{\mathbf{x}}$ *with* $\mathbf{x} = (x_1, \dots, x_n)$. *Then*

$$\sup_{t \le T} \|T_t^h F(\mu) - T_t F(\mu)\| \le hC(m, T, \|\mu\|) \|g^m\|_{C^2(X^m)}$$

with some constant C *depending on* $m, T, \|\mu\|$.

Proof. Notice first of all that (85) is applicable to B_1 of form (50), since it is well known (see e.g. [32]) that the fractional Laplacian $|\Delta|^\alpha$ has form (83) with $\nu(x, dy) = |y|^{-d-\alpha} dy$. Next, the standard representation for the difference of the action of two semigroups reads as

$$T_t F(\mu) - T_t^h(\mu) = \int_0^t \frac{d}{ds} T_{t-s}^h P_h T_s F(\mu)\, ds = \int_0^t U_{t-s}^h (P_h L_B - L_B^h P_h) T_s F(\mu)\, ds, \qquad (88)$$

where P_h denotes the projection on the functions depending on Dirac measures of the form $h\delta_{\mathbf{x}}$. Using (85), Theorem 4.2 (iv) and Proposition 5.3 (i) yields for $\|(P_h L_B - L_B^h P_h) T_s F(\mu)\|$ the estimate $hC(m, T, \|\mu\|) \|g^m\|_{C^2(X^m)}$, which directly implies the statement of the Theorem.

6 The central limit

We shall start with elementary results (Proposition 6.1 and 6.2 below) from the theory of Feller processes.

By the (time dependent) *generator* of a Feller backward propagator $U(t,s)$ in $C_\infty(X)$ we shall mean the operator

$$\Lambda_t g = \lim_{s\to t, s\le t} \frac{U(t,s)g-g}{t-s} = \lim_{r\to t, r\ge t} \frac{U(r,t)g-g}{r-t}$$

with the domain being the space of $g \in C_\infty(X)$ for which these two limits exist and coincide.

Proposition 6.1 *Suppose Ψ_t is the Feller semigroup of a Feller process Z_t on a Borel subset Z of X , and let Ω_t be a strongly continuous family of homeomorphisms of X (bijections continuous together with their inverses). Then the process $Y_t = \Omega_t(Z_t)$ in the family of subsets $\Omega_t(Z)$ of X is a nonhomogeneous Feller process whose (backward) propagator is given by*

$$U^Y(t,s)f(y_s) = \mathbf{E}^{t-s}_{\Omega_s^{-1}y_s}\Omega_t f = (\Omega_s^{-1})\Psi_{t-s}\Omega_t f(y_s), \tag{89}$$

where $\Omega_t f(y) = f(\Omega_t(y))$ and where $\mathbf{E}^{t-s}_z = \mathbf{E}^{t,s}_z$ denotes the expectation with respect to the probability distribution specified by the process Z_t starting at z.

Proof. Formula (89) follows from definition. One needs only to observe that $U^Y(t,s)$ is a strongly continuous family of contractions as it is a composition of such families.

Proposition 6.2 *Under the assumptions of Proposition 6.1 suppose that X is a topological linear space and that $\Omega_t(z) = (z-\xi_t)/a$, where a is a real constant and ξ_t, $t \ge 0$, is a differentiable curve in X . Suppose the domain of the generator L of Ψ_t contains a dense subspace D of $C_\infty(X)\cap C^2(X)$. Then the domain of the generator Λ_t of Y_t contains D at any time and is given there by the formula*

$$\Lambda_t f = \Omega_t^{-1} L \Omega_t f - \frac{1}{a}\left(\frac{\partial f}{\partial z}, \dot{\xi}_t\right). \tag{90}$$

Proof. Follows immediately by differentiating (89) using the product rule.

We can start now the analysis of the process of fluctuations around the evolution described by the kinetic equations.

Proposition 6.3 *Let $F(Y)$ be given by (57)) with $g = g^m \in C^{2,sym}_\infty(X^m)$. Let $\mu_t = \mu_t(B)$ solves the kinetic equation (10) and $\Lambda_t^{B,h}$ denote the generator of the Feller process $h^{-1/2}(Z^h_B(t)-\mu_t)$ (where $Z^h_B(t)$ is specified by the generator (79), (80)) given by (90) with $a = h^{1/2}$ and $L = L^h_B$. Then $\Lambda_t^{B_k,h}F$ is a polynomial functional of the form*

$$\Lambda_t^{B_k h}F(Y) = \sum_{l=2}^m h^{-1+l/2}\int_{X^{m-l}} \Phi_h^k[\tilde{B}_k^l g_{\mathbf{z}}](\sqrt{h}Y+\mu_t)Y^{\tilde{\otimes}(m-l)}(d\mathbf{z})$$
$$+ h^{-1/2}\int_{X^{m-1}} (\Phi_h^k[\tilde{B}_k^1 g_{\mathbf{z}}](\sqrt{h}Y+\mu_t) - (B_k g^+_{\mathbf{z}}, \mu_t^{\tilde{\otimes}k}))Y^{\tilde{\otimes}(m-1)}(d\mathbf{z}). \tag{91}$$

In particular,

$$\Lambda_t^{B_1,h}F(Y) = \int_{X^{m-l}} \int_X B_1 g_{\mathbf{z}}(y)(dy) Y^{\tilde{\otimes}(m-1)}(d\mathbf{z})$$
$$+ \sum_{l=2}^{m} h^{-1+l/2} \int_{X^{m-l}} \int_X \tilde{B}_1^l g_{\mathbf{z}}(y)(\sqrt{h}Y + \mu_t)(dy) Y^{\tilde{\otimes}(m-l)}(d\mathbf{z}). \tag{92}$$

Proof. Applying (77) to

$$\Omega_t F(Y) = F((Y-\mu_t)/\sqrt{h}) = h^{-m/2} \sum_{p=0}^{m} (-1)^{m-p} (g, Y^{\tilde{\otimes} p} \otimes \mu_t^{\tilde{\otimes}(m-p)})$$

yields

$$h^{k-1} L_k^h \Omega_t F(Y) = h^{-m/2} \sum_{p=1}^{m} (-1)^{m-p} \sum_{l=1}^{p} h^{l-1}$$
$$\times \int_{X^{m-p}} \int_{X^{p-l}} \left(\Phi_h^k [\tilde{B}^l g_{\mathbf{w},\mathbf{u}}](Y) \right) Y^{\tilde{\otimes}(p-l)}(d\mathbf{w}) \mu_t^{\tilde{\otimes}(m-p)}(d\mathbf{u}),$$

and consequently

$$h^{k-1} \Omega_t^{-1} L_k^h \Omega_t F(Y)) = h^{-m/2} \sum_{l=1}^{m} h^{l-1} \sum_{p=l}^{m} (-1)^{m-p}$$
$$\times \int_{X^{m-p}} \int_{X^{p-l}} \left(\Phi_h [\tilde{B}^l g_{\mathbf{w},\mathbf{u}}](\sqrt{h}Y + \mu_t) \right) (\sqrt{h}Y + \mu_t)^{\tilde{\otimes}(p-l)}(d\mathbf{w}) \mu_t^{\tilde{\otimes}(m-p)}(d\mathbf{u})$$
$$= h^{-m/2} \sum_{l=1}^{m} h^{l-1} \sum_{r=0}^{m-l} \left[\sum_{p=r+l}^{m} C_{m-l-r}^{m-p} (-1)^{m-p} \right]$$
$$\times \int_{X^r} \int_{X^{m-l-r}} \left(\Phi_h^k [\tilde{B}^l g_{\mathbf{z},\mathbf{y}}](\sqrt{h}Y + \mu_t) \right) h^{r/2} Y^{\tilde{\otimes} r}(d\mathbf{z}) \mu_t^{\tilde{\otimes}(m-l-r)}(d\mathbf{y}),$$

which yield all but the last terms on the r.h.s. of (91) by the obvious identity

$$\sum_{p=n}^{m} (-1)^{m-p} C_{m-n}^{m-p} = \begin{cases} 1, & n = m \\ 0, & n < m \end{cases}.$$

Subtracting

$$h^{-1/2} \left(\frac{\delta F}{\delta Y}(Y), \dot{\mu}_t \right) = h^{-1/2} \int \int (B_k g_{\mathbf{z}}^{+}(\mathbf{y}) Y^{\tilde{\otimes}(m-1)}(d\mathbf{z}) \mu_t^{\tilde{\otimes} k}(d\mathbf{y})$$

yields (91).

As the terms of order $h^{-1/2}$ cancel in (91) one obtains

Corollary 4

$$\Lambda_t^{B_k,h} F(Y) = \Lambda_t^{B_k} F(Y) + O(\sqrt{h}) \tag{93}$$

with

$$\Lambda_t^{B_k,h}F(Y) = \int_{X^{m-1}} \int_{X^k} (B_k(g^+_{\mathbf{w}}))(\mathbf{v})(Y \otimes \mu_t^{\tilde{\otimes}(k-1)})(d\mathbf{v})Y^{\tilde{\otimes}(m-1)}(d\mathbf{w})$$
$$+ \int_{X^{m-2}} \int_{X^k} (\tilde{B}_k^2(g^+_{\mathbf{w}}))(\mathbf{v})(\mu_t^{\tilde{\otimes}k})(d\mathbf{v})Y^{\tilde{\otimes}(m-2)}(d\mathbf{w}) \tag{94}$$

(the second term vanishes for $m = 1$), and $O(\sqrt{h})$ being a polynomial functional of Y of order $m + k - 1$ with coefficients bounded by the maximum of the sup-norms of the functions $\tilde{B}_k^l g, l = 1, \ldots, m$, uniformly for bounded μ_t.

The operators (94) are quite fundamental. Of course, they are extended by linearity to arbitrary polynomial functionals F. The next statement gives the alternative representations of these operators in terms of functional derivatives.

Corollary 5 *On functionals F of form (67) the action of the generators $\Lambda_t^{B_k}$ from (94) is given by*

$$\Lambda_t^B F(Y) = \left(\tilde{B}_k^1 \frac{\delta F}{\delta Y}, Y \otimes \mu_t^{\tilde{\otimes}(k-1)}\right) + \left(\tilde{B}_k^2 \frac{\delta^2 F}{\delta Y^2}, \mu_t^{\tilde{\otimes}k}\right), \tag{95}$$

or more explicitly

$$\Lambda_t^{B_k} F(Y) = \left(B_k \left(\frac{\delta F}{\delta Y}\right)^+, Y \otimes \mu_t^{\tilde{\otimes}(k-1)}\right)$$
$$+ \left(\frac{1}{2} B_k \sum_{i,j=1}^k \frac{\delta^2 F}{\delta Y(y_i)\delta Y(y_j)} - \left(B_k^{y_1,\ldots,y_k} \sum_{i,j=1}^k \frac{\delta^2 F}{\delta Y(z_i)\delta Y(y_j)}\right)|_{\forall i\ z_i=y_i}, \mu_t^{\tilde{\otimes}k}\right). \tag{96}$$

Formula (93) can be considered as a basis for an analytic study of the limiting fluctuation process, as it shows that the generator of the fluctuation process $h^{-1/2}(Z_B^h(t) - \mu_t)$ converges on polynomial functionals to the operator (94). What one needs here is, of course, a rigorous convergence result for the corresponding propagators, at least on some class of functionals. We shall demonstrate an analytic approach for obtaining such a result on our basic model of Section 4. Sticking to the tradition, we shall consider the linear functionals on measures, but will give the precise estimates for the remainder, which seem to be new even when applied to interacting diffusions (see [26], where already the uniform convergence with respect to the norm of a linear function and without our $\sqrt{h}$ estimate is presented as a significant progress compared to the usual result obtained by standard probabilistic method based on the compactness of approximating processes of fluctuation).

But first three general observations on the generator (94) are in order: (i) the propagator arising from (94) can be often easily constructed, because, as one sees from (95), this propagator preserves the space of polynomials of any given order; (ii) formula (95) indicates that the limiting process, whenever it is well defined, is an infinite dimensional Gaussian Orstein-Uhlenbeck process, but we shall not be concerned here with the question of existence of such a process or with its properties (however, see [45] and references therein for the general theory of such processes, and [22], [23] for the corresponding infinite dimensional system of Ito's stochastic equations arising in the context of interacting diffusions); (iii) in case when solutions to kinetic equations are regularizing in the sense that the solution measure is absolutely continuous with respect to Lebesgue measure for all positive times (as in our basic example of interacting stable processes), the norm of the measure $h^{-1/2}(Z_B^h(t)-\mu_t)$ is of order $h^{-1/2}$ for all times, and hence the family of such measures can not converge weakly to a measure. In other words, one can expect only "much weaker" convergence. On the level of propagators, one can expect convergence on polynomial functionals with only sufficiently smooth coefficients. Moreover, any measure with a smooth density on $\mathbf{R}^d$ can be approximated by the sum of the Dirac measures $h\delta_{\mathbf{x}}$ in such a way that $|h\delta_{\mathbf{x}}-\nu|h^{-1/2}$ is bounded or even convergent to zero in the norm of the space $(C^1(X))^*$. Hence the appearance of the expression $\|(Z_B^h(0)-\mu_0)/\sqrt{h}\|_{(C^1(X))^*}$ below.

We shall consider now our basic model of Section 4. As it follows from (94), under the assumptions of Theorem 5.3 the generator of the limiting process $\Lambda_t^{B_1}+\Lambda_t^{B_2}$ defines the evolution of linear functionals $F_g(\mu)=(g,\mu)$ by means of the equation

$$\dot{g}(x)=-\sigma(x)|\Delta|^{\alpha/2}g(x)+\int(V(x,y)\nabla g(x)+V(y,x)\nabla g(y))\mu_t(dy), \tag{97}$$

which in the inverse time gives the transformation $G_{t,s}:g_t\mapsto g_s$ that is bounded in $C_\infty(X)$ and in $C_\infty\cap C^2(X)$ (again according to Theorem 4.1 and its obvious modification).

To prove our central limit we shall need the following auxiliary result.

Lemma 6.1 *Under the assumptions of Theorem 5.3 let $g^2\in C_\infty^{sym}(X^2)\cap C^2(X^2)$. Then*

$$\sup_{t\le T}\mathbf{E}\left(g^2,\left(\frac{Z_B^h(t)-\mu_t(\mu_0)}{\sqrt{h}}\right)^{\tilde{\otimes}2}\right)$$
$$\le C(T,\|Z_B^h(0)\|)\|g^2\|_{C^2(X)}\left(1+\sqrt{h}\left\|\frac{Z_B^h(0)-\mu_0)}{\sqrt{h}}\right\|_{C^1(X))^*}+\left\|\frac{Z_B^h(0)-\mu_0)}{\sqrt{h}}\right\|^2_{(C^1(X))^*}\right) \tag{98}$$

with $C(T,\|\mu_0\|)$ being a constant not depending on h.

Proof. One has

$$\mathbf{E}\left(g^2,\left(\frac{Z_B^h(t)-\mu_t(Z_B^h(0))}{\sqrt{h}}\right)^{\tilde{\otimes}2}\right)$$
$$=\mathbf{E}\left(g^2,\left(\frac{Z_B^h(t)-\mu_t(Z_B^h(0))}{\sqrt{h}}\right)^{\tilde{\otimes}2}\right)+\left(g^2,\left(\frac{\mu_t,Z_B^h(0)-\mu_t(\mu_0)}{\sqrt{h}}\right)^{\tilde{\otimes}2}\right)$$
$$+2\mathbf{E}\left(g^2,\frac{Z_B^h(t)-\mu_t(Z_B^h(0))}{\sqrt{h}}\otimes\frac{\mu_t,Z_B^h(0)-\mu_t(\mu_0)}{\sqrt{h}}\right).$$

The estimate of the first term corresponds to the first term in the bracket on the r.h.s. of (98), which follows from Theorem 5.3 and the formula

$$\mathbf{E}\left(g^2,\left(\frac{Z_B^h(t)-\mu_t(Z_B^h(0))}{\sqrt{h}}\right)^{\tilde{\otimes}2}\right)=\frac{1}{2h}\mathbf{E}(g^2,Z_B^h(t)^{\otimes 2}-\mu_t(Z_B^h(0))^{\otimes 2}$$
$$+\mu_t(Z_B^h(0))\otimes(\mu_t(Z_B^h(0))-Z_B^h(t))+(\mu_t(Z_B^h(0))-Z_B^h(t))\otimes\mu_t(Z_B^h(0))).$$

The required estimate for the second and third terms are given by the third and second term respectively in the bracket on the r.h.s. of (98), which follows from Theorem 4.2 (v).

We can prove now the main result of this section.

Theorem 6.1 *Under the assumptions of Theorem 5.3 let* $g \in C^2_\infty(X)$*. Then*

$$\sup_{t\le T}\left|\mathbf{E}\left(g, \frac{Z^h_B(t)-\mu_t(\mu_0)}{\sqrt{h}}\right) - \left(G_{t,0}g, \frac{Z^h_B(0)-\mu_0}{\sqrt{h}}\right)\right|$$
$$\le C(T, \|Z^h_B(0)\|)\sqrt{h}\|g\|_{C^2(X)}\left(1+\left\|\frac{Z^h_B(0)-\mu_0}{\sqrt{h}}\right\|^2_{(C^1(X))^*}\right). \tag{99}$$

Proof. Let $U^{fluc}_h(t,s)$ be the backward propagator corresponding to the process $(Z^h_B(t)-\mu_t)/\sqrt{h}$. The l.h.s. of (98) can be written as

$$\sup_{t\le T}\left|(U^{fluc}_h(t,0)F_g(\xi_0) - (G_{t,0}g,\xi_0)\right|$$

with $\xi_0 = (Z^h_B(0)-\mu_0)/\sqrt{h}$, which equals

$$\sup_{t\le T}\int_0^t U^{fluc}_h(s,0)(\Lambda^{B,h}_s - \Lambda^B_s)G_{t,s}g\, ds(\xi_0).$$

As $(\Lambda^{B,h}_s - \Lambda^B_s)G_{t,s}g$ is a quadratic functional proportional to $\sqrt{h}$, the required estimate follows from Lemma 6.1.

7 Processes changing the number of particles

The methods developed above can be applied also to the analysis of processes changing the number of particles, as we are going to demonstrate now on the example of coagulation-fragmentation models.

To generalize Section 5 in a way to include processes changing the number of particles, assume that D and X are the same as in Theorem 2.1 and let $B_k : C^{sym}(\mathcal{X}) \mapsto C^{sym}(X^k)$, $k = 1,\dots,K$ be of the form (12) with $\mathcal{B}_k$ being conditionally positive operators in $C^{sym}(X^k)$ (that stand for the processes preserving the number of particles and that were denoted B_k in Sections 5 and 6) and P_k being a family of symmetric transition kernels from X^k to $\mathcal{X}$. A process $Z_B(t)$ of k-ary interaction in X (possibly changing the number of particles) can be defined again through the generator (72), which can be written in more detail as

$$L_Bf(\mathbf{x}) = \sum_{k=1}^K L_kf(\mathbf{x}) = \sum_{k=1}^K \sum_{I\subset\{1,\dots,l\}:|I|=k} [(\mathcal{B}^I_k f)(\mathbf{x}) + \int (f(\mathbf{y},\mathbf{x}_{\bar{I}}) - f(\mathbf{x}))P_k(\mathbf{x}, d\mathbf{y})] \tag{100}$$

for $\mathbf{x} = (x_1,\dots,x_l)$ with arbitrary l. The scaled process on $\mathcal{M}^+_{\delta,h}(X)$ is again given by (73), (74). Our calculations in Sections 5 and 6 were carried out in such a way that they are straightforwardly generalized to the new model of this section. For instance, (77), (78) still hold, where for the integral operator $B_k : C^{sym}(\mathcal{X}) \mapsto C^{sym}(X^k)$ of the form

$$B_k f(\mathbf{x}) = \int_{\mathcal{X}} (f(\mathbf{y}) - f(\mathbf{x}))P(\mathbf{x}, d\mathbf{y}), \qquad \mathbf{x} \in X^k,$$

the operator $\tilde{B}_k^l : C^{sym}(X^l) \mapsto C^{sym}(X^k)$ is given by

$$\tilde{B}_k^l g(\mathbf{x}) = \sum_{q=1}^{l} (-1)^{l-q} \int\int\int g(\mathbf{y}, \mathbf{z})(\delta_{\mathbf{u}} - \delta_{\mathbf{x}})^{\tilde{\otimes} q}(d\mathbf{y})P(\mathbf{x}, d\mathbf{u})\delta_{\mathbf{x}}^{\tilde{\otimes}(l-q)}(d\mathbf{z}). \tag{101}$$

This leads to the corresponding generalization of Theorems 5.1 with the law of large number being again specified by (10) and the limiting process of fluctuations by generators (94).

Remark. Our decomposition of B_k from (1.12) into the sum of $\mathcal{B}_k$ (the number of particles preserving part) and the remaining jump-type part is not unique, as the jumps preserving the number of particles (like in Boltzmann collisions) can be put in either of these two parts.

In case of pure jump interactions $\mathcal{B}_k = 0$ in (100) and kinetic equation (10) takes the form

$$\frac{d}{dt}(g, \mu_t) = \sum_{k=1}^{K} \int\int (g^+(\mathbf{y}) - g^+(x_1, \ldots, x_k))P_k(x_1, \ldots, x_k; d\mathbf{y})\mu_t^{\tilde{\otimes} k}(dx). \tag{102}$$

Hence the generator (94) of the limiting process for fluctuation can be written as

$$\begin{aligned} \Lambda_t^{B_k} F(Y) &= \int\int (\delta F(\mathbf{z}) - \delta F(\mathbf{y}))P_k(\mathbf{y}, d\mathbf{z})(Y \otimes \mu_t^{\tilde{\otimes} k-1)})(d\mathbf{y}) \\ &+ \frac{1}{2} \int\int (\delta^2 F(\mathbf{z}, \mathbf{z}) + \delta^2 F(\mathbf{y}, \mathbf{y}) - 2\delta^2 F(\mathbf{y}, \mathbf{z}))P_k(\mathbf{y}, d\mathbf{z})\mu_t^{\tilde{\otimes} k}(d\mathbf{y}) \end{aligned} \tag{103}$$

and the dynamics of the invariant linear functions $F_g(Y) = (g, Y)$ is given by the equation

$$\dot{g}(x) = \sum_{k=1}^{K} \sum_{n=1}^{N} \int \left[\sum_{i=1}^{n} g(z_i) - g(x) - \sum_{i=1}^{k-1} g(y_i) \right] P_k(x, y_1, \ldots, y_{k-1}; dz_1 \cdots dz_n)\mu_t^{\tilde{\otimes}(k-1)}(d\mathbf{y}). \tag{104}$$

As in the previous section, we shall denote by $G_{t,s}$ the corresponding evolution operators on g in inverse time (backward propagator).

We shall consider now the standard model of coagulation and fragmentation combined with possible collision breakage, where $X = \mathbf{R}_+$, P_k does not vanish only for k being two and one, so that

$$P_2(x_1, x_2; d\mathbf{y}) = K(x_1, x_2)\delta(x_1 + x_2 - y)dy + C(x_1, x_2, y_1)\delta(x_1 + x_2 - y_1 - y_2)dy_1dy_2,$$

and $P_1(x; dy_1dy_2) = F(x, y_1)\delta(y_1 + y_2 - x)\, dy_1dy_2$. The continuous functions K, C, F are called the coagulation, collision and fragmentation kernels respectively. The corresponding law of large numbers (102) takes the form

$$\frac{d}{dt}(g,\mu_t) = \int_0^\infty \int_0^\infty (g(x_1+x_2) - g(x_1) - g(x_2))K(x_1,x_2)\mu_t^{\tilde{\otimes}2}(dx_1dx_2)$$
$$+ \int_0^\infty \int_0^\infty \int_0^{x_1+x_2} dz(g(z) + g(x_1+x_2-z) - g(x_1) - g(x_2))C(x_1,x_2,z)\mu_t^{\tilde{\otimes}2}(dx_1dx_2)$$
$$+ \int_0^\infty \int_0^x dz(g(z) + g(x-z) - g(x))F(x,z)\mu_t(dx) \quad (105)$$

(called Smoluchovski's equation in case of vanishing C and F). Assume for simplicity that all intensities are bounded (more realistic assumptions will be discussed in [41]):

$$\sup_{x_1,x_2,x} \max(K(x_1,x_2), \int C(x_1,x_2,z)dz, \int F(x,z)dz) < \infty. \quad (106)$$

Then it is well known that (105) is well posed (see e.g. [52] for coagulations and [35] in general case), i.e. for any finite μ_0 with a finite second moment $\int x^2\mu_0(dx) < \infty$ there exists a unique bounded solution μ_t of (105) with a bounded second moment and preserving the mass, i.e. such that $\int x\mu_0(dx) = \int x\mu_t(dx)$. In [35] and [39] it is shown that this equation holds also in the strong sense with the derivative being understood in the sense of the norm of $\mathcal{M}(X)$. This allows to differentiate this equation with respect to the initial measure μ_0. Equation for all derivatives are obviously linear and by a simple induction one sees that

$$\sup_{x_1,\dots,x_l} \left\| \frac{\delta^l \mu_t}{\delta \mu_0^l}(x_1,\dots,x_l) \right\| \le C^l(\mu_0)t^l l! \quad (107)$$

for all l with some constant $C(\mu_0)$.

Similar to Theorem 5.3 we get now the following result on the rate of convergence to the law of large numbers (105).

Theorem 7.1 *Assume* (106) *in the coagulation-collision-fragmentation model* (105). *Let* $g^m \in C_\infty(X^m)$ *and* $F(\mu) = (g^m, \mu^{\tilde{\otimes}m})$. *Assume* $\mu = h\delta_{\mathbf{x}}$ *with* $\mathbf{x} = (x_1,\dots,x_n)$. *Then*

$$\sup_{t\le T} \|T_t^h F(\mu) - T_t F(\mu)\| \le hC(m,T,\|\mu\|)\|g\|_{C(X^m)}$$

with some constant C *depending on* $m, T, \|\mu\|$.

Proof. We shall again use (88). $T_sF(\mu)$ is an infinitely differentiable function, the bounds for the derivatives being given by (107). From (101) it follows that

$$\|\tilde{B}_k^l g\| \le 5^l C\|g\|/l!, \qquad k = 1,2,$$

and hence series (78) turns out to be convergent for $F_t(\mu) = F(\mu_t)$ and the sum over $l \ge 2$ is estimated by

$$\frac{1}{k!}\sum_{l=2}^\infty h^{l-1} \left\| \tilde{B}_k^l \frac{\delta^l F_t}{\delta \mu^l} \right\| \|h\delta_{\mathbf{x}}\|^k \le \frac{h}{k!}\sum_{l=2}^\infty h^{l-2}(C(\mu_0)t)^l \|h\delta_{\mathbf{x}}\|^k, \qquad k = 1,2,$$

so that this is convergent and of order $O(h)$ bounded for each t (and sufficiently small h). Hence $(P_h L_B - L_B^h P_h) T_s F(\mu)$ is of order h in (88), which implies the statement of the Theorem.

We can now obtain the central limit result for this model basically copying the arguments of the previous Section, even with additional simplifications that we do not need any smoothness of the coefficients of polynomial functions (since all B_k are bounded). Thus instead of Lemma 6.1 we get

Lemma 7.1 *Under the assumptions of Theorem 7.1 let $g^2 \in C_\infty^{sym}(X^2)$. Then*

$$\sup_{t\le T} \mathbf{E}\left(g^2, \left(\frac{Z_B^h(t)-\mu_t(\mu_0)}{\sqrt{h}}\right)^{\tilde\otimes 2}\right) \le C(T, \|Z_B^h(0)\|)\|g^2\| \left(1 + \left\|\frac{Z_B^h(0)-\mu_0}{\sqrt{h}}\right\|^2_{\mathcal{M}(X)}\right)$$

with $C(T, \|\mu_0\|)$ being a constant not depending on h.

And then similar modification of Theorem 6.1 yields the following central limit result.

Theorem 7.2 *Under the assumptions of Theorem 5.3 let $g \in C_\infty(X)$. Then*

$$\sup_{t\le T}\left| \mathbf{E}\left(g, \frac{Z_B^h(t)-\mu_t(\mu_0)}{\sqrt{h}}\right) - \left(G_{t,0}g, \frac{Z_B^h(0)-\mu_0}{\sqrt{h}}\right)\right|$$
$$\le C(T, \|Z_B^h(0)\|)\sqrt{h}\|g\| \left(1 + \left\|\frac{Z_B^h(0)-\mu_0}{\sqrt{h}}\right\|^2_{\mathcal{M}(X)}\right). \quad (108)$$

This result is new even for coagulation model (with vanishing C and F) and even without the estimate of convergence. The only previous result on the central limit for coagulation model appeared in [18], and it is devoted to the case of only discrete mass distribution. The arguments in [18] are quite different from ours and they do not give any estimates for convergence.

A Approximation of infinite-dimensional functions

Let B and B^* be a real separable Banach space and its dual with duality denoted by $(\cdot,\cdot)$ and the unit balls denoted by B_1 and B_1^*. It follows from the Stone-Weierstrass theorem that finite dimensional (or cylindrical) functions of form $F_f(v) = f((g_1, v), \ldots, (g_m, v))$ with $g_1, \ldots, g_m \in B$ and $f \in C(\mathbf{R}^m)$ are dense in the space $C(B_1^*)$ of $*$-weakly continuous bounded functions on the unit ball in B^* . We need a more precise statement that these approximations can be chosen in such a way that they respect differentiation. This fact is crucial for our exposition and the author did not find it in the literature. We shall need it only for the case of $B = C_\infty(X)$ with X being $\mathbf{R}^n$ or its submanifold, and we reduce attention only to this case.

We shall say that a family $P_1, P_2, \cdots$ of the linear contractions in B of the form

$$P_j v = \sum_{l=1}^{L_j} (w_j^l, v)\phi_j^l, \quad (109)$$

where ϕ_j^l and w_j^l are some finite linear independent sets from the unit balls B_1^* andB_1 respectively, form an approximative identity, if the sequence P_j converges strongly to the identity operator as $j \to \infty$.

To see the existence of such a family, let us choose a finite $\frac{1}{j}$-net $x_1, x_2, \dots, x_{L_j}$ in the ball $\{\|x\| \le j\}$, and let ϕ_j^l be a collection of continuous non-negative functions such that $\phi_j(x) = \sum_l \phi_j^l(x)$ belongs to $[0,1]$ everywhere, equals one for $\|x\| \le j$ and vanishes for $\|x\| \ge j+1$ and such that each ϕ_j^l equals one in a neighborhood of x_j^l and vanishes for $\|x - x_j^l\| \ge 2/j$. Then the operators

$$P_j f(x) = \sum_{l=1}^{L_j} f(x_j^l)\phi_j^l(x) = \sum_{l=1}^{L_j} (f, \delta_{x_j^l})\phi_j^l(x)$$

form an approximative identity in $B = C_\infty(X)$.

Proposition A.1 *Suppose a family $P_1, P_2, \cdots$ of finite dimensional linear contractions in B given by* (109) *form an approximative identity in B. Then*

(i) for any $F \in C(B_1^*)$ the family $F_j = F(P_j^*)$ converges to F uniformly (i.e. in the norm topology of $C(B_1^*)$),

(ii) there exist positive numbers $\epsilon_1, \epsilon_2, \cdots$ and a family of contractions Π_j on $C(B_1^*)$ with the range consisting of finite linear combinations of the analytic functions of $\mu \in B^*$ of the form

$$F_{j,\nu}(\mu) = \exp\left\{ -\epsilon_j \sum_{l=1}^{L_j} (\phi_j^l, \mu - \nu)^2 \right\}, \qquad \nu \in B^*,$$

such that $\Pi_j(F)$ converges to F uniformly on B_1^*,

(iii) if F is k times continuously differentiable in the sense that $\delta^k F(\mu)(v_1, \dots, v_k)$ exists and is a $*$-weakly continuous function of $k+1$ variables, then the derivatives of the order k of $\Pi_j(F)$ converge to the corresponding derivatives of F uniformly on B_1^*.

Proof. (i) Notice that

$$P_j^*(\mu) = \sum_{l=1}^{L_j} (\phi_j^l, \mu) w_j^l.$$

The required convergence for the functions of the form $F_g(\mu) = \exp\{(g, \mu)\}$, $g \in C_\infty(X)$ follows from the definition of the approximative identity, for

$$F_g(P_j^*(\mu)) = \exp\{(P_j g, \mu)\}.$$

For arbitrary $F \in C(B_1^*)$ the statement is obtained through its approximation by the linear combinations of exponential functions F_g (which is possible by the Stone-Weierstrass theorem).

(ii) For any j the functional $F_j(\mu) = F(P_j^*(\mu))$ clearly can be written in the form $F_j(\mu) = f_j(y(\mu))$ with $y(\mu) = \{(\phi_j^1, \mu), \dots, (\phi_j^{L_j}, \mu)\}$ and f_j being bounded continuous functions of L_j variables. Approximating f_j first by the functions $e^{\epsilon \Delta} f_j$ and the latter function by a linear combination h_j of the Gaussian functions of the type $\exp\{-\epsilon_j \sum_{l=1}^{L_j} (y_j - \xi_j)^2\}$, we then define $\Pi_j(F(\mu)) = h_j(y(\mu))$, which enjoys the required property.

(iii) One only needs to notice that if F is k times continuously differentiable, then

$$\delta^k F_j(\mu)(v_1,\ldots,v_k) = \delta^k F(P_j^*(\mu))(P_j^* v_1,\ldots,P_j^* v_k)$$

and then the result follows from (i).

B A combinatorial lemma

Proposition B.1 *For any natural k , there exists a linear mapping $f \mapsto \Phi^k[f] = \Phi[f]$ from the space $C^{sym}(X^k)$ to the polynomial functionals on measures of order k such that*

$$\sum_{I\subset\{1,\ldots,n\},|I|=k} f(\mathbf{x}_I) = \Phi^k[f](\delta_{\mathbf{x}}) \tag{110}$$

for an arbitrary point $x = (x_1,\ldots,x_n)$ in $\mathcal{X}$.

This mapping has the form

$$\Phi^k[f](Y) = (f, Y^{\tilde\otimes k}) + \sum_{l=1}^{k-1}(-1)^l(\Phi_l^k[f], Y^{\tilde\otimes(k-l)}) \tag{111}$$

with $\Phi_l^k[f]$being positivity preserving bounded linear operators from $C^{sym}(X^k)$ to the continuous functions of $k-l$ variables.

In particular,

$$\Phi^2[f](Y) = \frac{1}{2}\int\int f(y_1,y_2)Y(dy_1)Y(dy_2) - \frac{1}{2}f(y,y)Y(dy). \tag{112}$$

Remark. Both sides of (110) vanish in case $n < k$.

Proof. Observing that

$$\sum_{I\subset\{1,\ldots,n\},|I|=k} f(\mathbf{x}_1) = (f, (\delta_{\mathbf{x}})^{\tilde\otimes k}) - \sigma,$$

where σ denotes the sum over the combinations with not more than $k-1$ different points, one easily obtains that

$$\sum_{I\subset\{1,\ldots,n\},|I|=k} f(x_1) = (f, (\delta_{\mathbf{x}})^{\tilde\otimes k}) - \frac{1}{k!}\sum_{j=1}^{k-1} jj! \sum_{J\subset\{1,\ldots,n\},|J|=j} P^j f(\mathbf{x}_J,\mathbf{z})\delta_{\mathbf{x}}^{\tilde\otimes(k-j-l)}(d\mathbf{z}),$$

where P_j means the symmetrization over the first j variables of the gluing operator

$$Pf(x_1,\ldots,x_{n-1}) = f(x_1,x_1,x_2,\ldots,x_{n-1}).$$

From this formula the required statement follows by straightforward induction together with the explicit formula

$$\Phi_l^k(f) = \sum_{i_l=l}^{k-1} i_l \sum_{i_{l-1}=l-1}^{i_l-1} i_{l-1}\cdots i_2 \sum_{i_1=1}^{i_2-1} i_1 P^{i_1}\cdots P^{i_1} f.$$

As we are interested in the scaled transformations, we shall introduce the natural scaling in $\Phi[f]$ defining

$$\Phi_l^k[f](Y) = h^k \Phi^k[f](Y/h). \tag{113}$$

In this notations (110) yields the following simple but fundamental formula

$$h^k \sum_{I\subset\{1,\dots,n\},|I|=k} f(\mathbf{x}_I) = \Phi_h^k[f](h\delta_{\mathbf{x}}) = (f,(h\delta_x)^{\tilde{\otimes}k}) + \sum_{l=1}^{k-1}(-h)^l(\Phi_l^k[f],(h\delta_{\mathbf{x}})^{\tilde{\otimes}(k-l)}). \tag{114}$$

Acknowledgments. The author is thankful to Roberto Quezada Batalla for his hospitality in the Universidad Autonoma Metropolitana (UAM, Mexico city), this work being initiated during the author's visit to UAM in autumn 2004. The author is grateful to S. Gaubert and M. Akian for their hospitality in INRIA (France), this work being completed during the author's visit to INRIA in the spring 2006. The author thanks the referees for their constructive critique of the first draft of this paper and Martine Verneuille for quick professional typing of this manuscript in Latex.

References

[1] S. Albeverio, A. Hilbert, V. Kolokoltsov. Sur le comportement asymptotique du noyau associé une diffusion dégénéré (French) [Asymptotic behavior of the kernel associated with a degenerate diffusion]. C. R. Math. Acad. Sci. Soc. R. Can. **22:4** (2000), 151-159.

[2] D.J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. Bernoulli **5:1** (1999), 3-48.

[3] D. Applebaum. Lévy processes and stochastic calculus. Cambridge Studies in Advanced Mathematics, 93. Cambridge University Press, Cambridge, 2004.

[4] O.E. Barndorff-Nielsen et al (Eds.) Lévy processes: Theory and Applications. Birkhäuser, 2001.

[5] R. Bass. Stochastic differential equations with jumps. Probab. Surv. 1 (2004), 1-19 (electronic).

[6] V.P. Belavkin. Quantum branching processes and nonlinear dynamics of multiquantum systems. Dokl. Acad. Nauk SSSR **301:6** (1988), 1348-1352. Engl. Tansl. in Sov. Phys. Dokl. **33:8** (1988), 581-582. 47

[7] V. Belavkin, V. Kolokoltsov. On general kinetic equation for many particle systems with interaction, fragmentation and coagulation. Proc. R. Soc. Lond. A **459** (2002), 1-22.

[8] V. Bening et al. Fractionally stable distributions. In: V.Yu. Korolev, N.N Skvortsova (Eds.) Stochastic Models of Structural Turbulence in Plasma. Moscow State University, Moscow 2003 (in Russian).

[9] J.-M. Bony, Ph. Courrège, P. Priouret. Semi-groupes de Feller sur une variété à bord compacte et problèmes aux limites intégro-différentielle du second ordre donnant lieu au principe du maximum. Ann. Inst. Fourier, Grenoble **18:2** (1968), 369-521.

[10] W. Braun, K. Hepp. The Vlasov dynamics and its fluctuations in the 1/N limit of interacting particles. Comm. Math. Phys. **56** (1977), 101-113.

[11] R.A. Carmona, L. Xu. Diffusive hydrodynamic limits for systems of interacting diffusions with finite range random interaction. Comm. Math. Phys. **188:3** (1997), 565-584.

[12] Ph. Courr'ege. Sur la form integro-différentiélle du générateur infinitésimal d'un semigroupe de Feller sur une varit. Sém. Théori du Potentiel, 1965-66, Exposé 3.

[13] D. Dawson. Critical dynamics and fluctuations for a mean field model of cooperative behavior. J. Stat. Phys. **31** (1983), 29-85.

[14] D. Dawson. Measure-Valued Markov Processes. Hennequin P.L. (ed.) Ecole d'Eté de probabilités de Saint-Flour XXI-1991. Springer Lect. Notes Math. 1541 (1993), 1-260.

[15] D. Dawson et al. Generalized Mehler Semigroups and Catalytic Branching Processes with Immigration. Potential Anal. **21:1** (2004), 75-97.

[16] D.A. Dawson, J. Gärtner. Long time behavior of interacting diffusions. In: J.R. a Norris (Ed.) Stochastic Calculus in Application. Pitman Research Notes in Mathematics Series v. 197 (1988), 29-54.

[17] D.A. Dawson, J. Gärtner. Multilevel large deviations and interacting diffusions. Probab. Theory Relat. Fields **98:4** (1994), 423487.

[18] M. Deaconu, N. Fournier, E. Tanré. Rate of convergence of a stochastic particle e system for the Smoluchowski coagulation equation. Methodol. Comput. Appl. Probab. **5:2** (2003), 131-158.

[19] R.L. Dobrushin. Vlasov equations. Funct. Anal. Appl. **13** (1979), 115.

[20] R. Ferland, X. Fernique, G.Giroux. Compactness of the fluctuations associated with some generalized nonlinear Boltzmann equations. Canad. J. Math. **44:6** (1992), 1192-1205.

[21] S.N. Ethier, T.G. Kurtz. Markov Processes. Characterization and convergence. John Wiley Sons, 1986.

[22] B. Fernandez, S. Méléard. A Hilbertian approach for fluctuations on the McKean-Vlasov model. Stochastic Processes Appl. **71** (1997), 33-53.

[23] B. Fernandez, S. Méléard. Asymptotic behavior for interacting diffusion processes with space-time random birth. Bernoulli **6:1** (2000), 91-111. 48

[24] J. Gärtner. On the McKean-Vlasov limit for interacting diffusions. Math. Nachr. **137** (1988), 197-248.

[25] C. Graham, S. Méléard. Stochastic particle approximations for generalized ee Boltzmann models and convergence estimates. Ann. Prob. **25:1** (1997), 115-132.

[26] E. Giné, J.A. Wellner. Uniform convergence in some limit theorem for multiple particle systems. Stochastic Processes and their Applications **72** (1997), 47-72.

[27] F. Guias. Convergence properties of a stochastic model for coagulation-fragmentation processes with diffusion. Stoch. Anal. Appl. **19:2** (2001), 245-278.

[28] N. Jacob. Pseudo-differential operators and Markov processes. Vol. II: Generators and their potential theory. London: Imperial College Press, 2002.

[29] J. Jacod, A. Shiryaev. Limit theorems for stochastic processes. Second edition. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 288. Springer-Verlag, Berlin, 2003.

[30] M. Kac. Probability and related topics in physical science. Wiley Interscience, 1959.

[31] A.N. Kochubei. Parabolic pseudo-differentiable equation, supersingular integrals and Markov processes. Izvestia Akad. Nauk, Ser. Matem. **52:5** (1988), 909-934 (Russian), Math.-USSR Izv. **33:2** (1989), 233-259 (English).

[32] V. N. Kolokoltsov. Symmetric Stable Laws and Stable-Like Jump-Diffusions. Proc. London Math. Soc. **3:80** (2000), 725-768.

[33] V. N. Kolokoltsov. Semiclassical Analysis for Diffusions and Stichastic Processes. Springer Lecture Notes in Math. v. 1724, 2000.

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

[35] V. N. Kolokoltsov. On Extension of Mollified Boltzmann and Smoluchovski Equations to Particle Systems with a k-ary Interaction. Russian Journal of Math.Phys. **10:3** (2003), 268-295.

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

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

[38] V. N. Kolokoltsov. Hydrodynamic Limit of Coagulation-Fragmentation Type Models of k-nary Interacting Particles. Journal of Statistical Physics **115**, 5/6 (2004), 1621-1653.

[39] V. N. Kolokoltsov. Kinetic equations for the pure jump models of k-nary interacting particle systems. Markov Processes and Related Fields **12:1** (2006), 95-138.

[40] V. N. Kolokoltsov. On the regularity of solutions to the spatially homogeneous Boltzmann equation with polynomially growing collision kernel. Preprint Universidad Autonoma Metropolitana, 04.0402.1.I.003.2005, Mexico. Published in Adv.Stud.Cont.Math. **12:1** (2006).

[41] V.N. Kolokoltsov. The central limit theorem for the Smoluchowski coagulation model. In preparation.

[42] A. Kolodko, K. Sabelfeld, W. Wagner. A stochastic method for solving Smoluchowski's coagulation equation. Math. Comput. Simulation **49** (1999), 57-79.

[43] T. Komatsu. On the martingale problem for generators of stable processes with perturbations. Osaka J. Math. **21** (1984), 113-132.

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

[45] P. Lescot, M. Roeckner. Perturbations of Generalized Mehler Semigroups and Applications to Stochastic Heat Equation with Lévy Noise and Singular Drift. Potential Anal. **20:4** (2004), 317-344.

[46] F. Leyvraz. Scaling theory and exactly solved models in the kinetics of irreversible aggregation. Physics Reports **383**, 2-3 (2003), 95-212.

[47] V.P. Maslov. Perturbation Theory and Asymptotical Methods. Moscow State University Press, 1965 (in Russian). French Transl. Dunod, Paris, 1972.

[48] V.P. Maslov. Complex Markov Chains and Functional Feynman Integral. Moscow, Nauka, 1976 (in Russian).

[49] H.P. McKean. A class of Markov processes associated with nonlinear parabolic equations. Proc. Nat. Acad. Sci. **56** (1966), 1907-1911.

[50] [S. Méléard. Convergence of the fluctuations for interacting diffusions with jumps associated with Boltzmann equations. Stochastics Stochastics Rep. **63: 3-4** (1998), 195-225.

[51] J.M. van Neeren. Continuity and Representation of Gaussian Mehler Semigroups. Potential Anal. **13:3** (2000), 199-211.

[52] J. Norris. Cluster Coagulation. Comm. Math. Phys. **209**(2000), 407-435.

[53] R.L. Schilling. Growth and Hölder conditions for the sample paths of Feller processes. Probab. Theory Relat. Fields **112** (1998), 565-611.

[54] D. Stroock. Markov Processes from K. Itô's Perspective. Annals of Math. Studies 155, Princeton University Press, 2003.

[55] A.-S. Sznitman. Nonlinear Reflecting Diffusion Process and the Propagation of Chaos and Fluctuation Associated. Journ. Funct. Anal. **56** (1984), 311-336.

[56] A.-S. Sznitman. Topics in Propagation of Chaos. In: Ecole d'Eté de Probabilités de Saint-Flour XIX-1989. Springer Lecture Notes Math. 1464 (1991), 167-255.

[57] H. Tanaka, M. Hitsuda. Central limit theorems for a simple diffusion model of interacting particles. Hiroshima Math. J. **11** (1981), 415-423

[58] V.V. Uchaikin, V.M. Zolotarev. Chance and Stability: Stable Distributions and their Applications. VSP, 1999.

[59] K. Uchiyama. Scaling limit of interacting diffusions with arbitrary initial distributions. Prob. Theory Relat. Fields **99** (1994), 97-110.