THE CENTRAL LIMIT THEOREM FOR THE SMOLUCHOVSKI COAGULATION MODEL *

Vassili N. Kolokoltsov[†]

Abstract

The general model of coagulation is considered. For basic classes of unbounded coagulation kernels the central limit theorem (CLT) is obtained for the fluctuations around the dynamic law of large numbers (LLN) described by the Smoluchovski equation. A rather precise rate of convergence is given both for LLN and CLT.

1 Introduction

In this paper we shall establish a central limit theorem for the fluctuations of the Markus–Lushnikov process in the limit of large particle numbers. Consider, for now, the case where coagulation rates are a function only of particle masses, so that any two particles, of masses x and x' say, coagulate to form a particle of mass $x+x'$ at a given rate $hK(x,x')$. Here, $N=1/h$ is the number of initial particles and the coagulation rates are scaled to give the following law of large numbers. The process of empirical particle distributions $Z^h=(Z^h_t)_{t\geq 0}$ converges weakly, as $h\to 0$ (or equivalently $N\to\infty$), when Z^h_0 converges weakly to μ_0, to the solution $(\mu_t)_{t\geq 0}$ of Smoluchowski's coagulation equation

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

Here, for suitable measures μ and μ' on $(0,\infty)$, $K(\mu,\mu')$ is the signed measure, given by

$$(f,K(\mu,\mu'))=\int_{(0,\infty)^2}(f(x+x')-f(x)-f(x'))K(x,x')\mu(dx)\mu'(dx'),$$

for suitable measurable functions f. Our result concerns the limiting distribution of the fluctuations

$$F^h_t=(Z^h_t-\mu_t)/\sqrt{h}.$$

This is of interest if we consider the Markus–Lushnikov model as representing a good mathematical description in some applied context and wish to understand, for large particle numbers,

*arXiv:0708.0329v1[math.PR], the results were presented on the Obervolfach conference (September 2007); final version published in PTRF (online 2008)

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

how this model deviates from the deterministic evolution given by Smoluchowski's equation. It is also important in quantifying the stochastic errors which may arise in a computational approach to Smoluchowski's equation using Monte–Carlo techniques. Though formal calculations leading to the formal expression of the covariance of the limiting infinite dimensional Gaussian (Ornstein-Uhlenbeck) process are not difficult, the rigorous identification of the limit turns out to be not simple, this problem being placed as problem 10 in the list of open mathematical problems on the coagulation theory in the well known review [1].

To make things more precise, we shall fix some notations. We shall denote by X a locally compact topological space equipped with its Borel sigma algebra and by E a given continuous non-negative function on X such that $E(x) \to \infty$ as $x \to \infty$. Denoting by X^0 a one-point space and by X^j the powers $X \times \ldots \times X (j\text{-times})$ considered with their product topologies, we shall denote by $\mathcal{X}$ their disjoint union $\mathcal{X} = \bigcup_{j=0}^{\infty} X^j$, which is again a locally compact space. In applications, X specifies the state space of a single particle, $\mathcal{X}$ stands for the state space of a random number of similar particles, and E describes some key parameter of a particle. In the standard model $X = \mathbf{R}_+ = \{x > 0\}$ and $E(x) = x$ denotes the mass of a particle.

By $C(X)$ (respectively $C_\infty(X)$) we denote the Banach space of continuous bounded functions on X (respectively its subspace of functions vanishing at infinity) with the sup-norm denoted by $\|\cdot\|$, by $\mathcal{M}(X)$ - the Banach space of finite Borel measures on X with the norm also denoted by $\|\cdot\|$, and by $\mathcal{M}^+(X)$ - the set of its positive elements. The brackets (f, Y) denote the usual pairing (given by the integration) between functions f and measures Y, and $|\mu|$ for a signed measure μ denotes its total variation measure. The elements of $\mathcal{X}$ will be denoted by bold letters, e.g. $\mathbf{x} = (x_1, \ldots, x_n) \in X^n \subset \mathcal{X}$. For a subset I in $\{1, \ldots, n\}$ we shall denote by $|I|$ and $\bar{I}$ respectively its cardinality and its complement in $\{1, \ldots, n\}$, and by $\mathbf{x}_I$ the element of $X^{|I|}$ given by the collection of x_i with $i \in I$.

Assume that we are given a continuous transition kernel $K(x_1, x_2; dy)$ from $X \times X$ to X, i.e. a continuous function from $X \times X$ to $\mathcal{M}^+(X)$ (the latter equipped with its weak topology). This kernel will be called the coagulation kernel and it will be assumed to preserve E, i.e. $K(x_1, x_2; dy)$ has support contained in the set $\{y : E(y) = E(x_1) + E(x_2)\}$. Moreover, $K(x_1, x_2; dy)$ is symmetric with respect to permutation of x_1 and x_2 and has intensity $K(x_1, x_2) = \int_X K(x_1, x_2; dy)$ enjoying the following additive upper bound:

$$K(x_1, x_2) \leq C(1 + E(x_1) + E(x_2)) \tag{1.1}$$

with some constant $C > 0$ and all x_1, x_2.

The process of coagulation that we are going to analyze here is a Markov process $Z(t)$ on $\mathcal{X}$ specified by the generator

$$Lg(\mathbf{x}) = \sum_{I \subset \{1, \ldots, n\} : |I| = 2} \int (g(\mathbf{x}_{\bar{I}}, y) - g(\mathbf{x})) K(\mathbf{x}_I; dy) \tag{1.2}$$

(where $\mathbf{x} = (x_1, \ldots, x_n)$) of its Markov semigroup acting on an appropriate space of functions on $\mathcal{X}$. It is known and not difficult to deduce from the theory of jump type processes (see e.g. [5]) that the process $Z(t)$ is well defined by this generator (see e.g. a detailed probabilistic description of $Z(t)$ in [34]). In the next Section the analytic properties of the Markov semigroup specified by L will be made precise.

The transformation

$$\mathbf{x} = (x_1, \dots, x_n) \mapsto h\delta_{\mathbf{x}} = h(\delta_{x_1} + \cdots \delta_{x_n}), \tag{1.3}$$

with h being a positive (scaling) parameter, maps $\mathcal{X}$ to the space $\mathcal{M}_{h\delta}(X)$ of positive measures on X of the form $h\delta_{\mathbf{x}}$. By Z_t^h we shall denote a Markov process on $\mathcal{M}_{h\delta}(X)$ obtained from $Z(t)$ by transformation (1.3) combined with the scaling of L by h, i.e. Z_t^h is defined through the generator

$$\begin{aligned} L_h G_g(h\delta_{\mathbf{x}}) &= h \sum_{I\subset\{1,\dots,n\}:|I|=2} \int (g(\mathbf{x}_{\bar{I}}, y) - g(\mathbf{x}))K(\mathbf{x}_I; dy) \\ &= h \sum_{I\subset\{1,\dots,n\}:|I|=2} \int (G_g(h\delta_{\mathbf{x}} + h(\delta_y - \delta_{\mathbf{x}_I})) - G_g(h\delta_{\mathbf{x}}))K(\mathbf{x}_I; dy) \end{aligned} \tag{1.4}$$

on $C(\mathcal{M}_{h\delta}(X))$, where $G_g(h\delta_{\mathbf{y}}) = g(\mathbf{y})$ for any $\mathbf{y} \in \mathcal{X}$.

The law of large numbers dynamics (LLN) for the processes Z_t^h is given by the kinetic equation, whose most natural form is the weak one, i.e. it is the equation

$$\frac{d}{dt}(g, \mu_t) = \frac{1}{2} \int_{X\times X} \int_X (g(y) - g(x_1) - g(x_2))K(x_1, x_2; dy)\mu_t(dx_1)\mu_t(dx_2) \tag{1.5}$$

on μ_t that has to hold for all $g \in C_\infty(X)$. It is known (see [34]) that if a family of initial measures $h\delta_{\mathbf{x}(h)}$ for Z_t^h is uniformly bounded with bounded moments of order $\beta \geq 2$, i.e. if

$$\sup_h \int_X (1 + E^\beta(y))h\delta_{\mathbf{x}(h)}(dy) < \infty, \tag{1.6}$$

and if $h\delta_{\mathbf{x}(h)}$ tends $*$-weakly to a measure μ_0 on X, as $h \to 0$, then the process Z_t^h with the initial data $h\delta_{\mathbf{x}(h)}$ tends weakly to a bounded solution μ_t of (1.5) with initial condition μ_0 that preserves E and has bounded moments of order β, i.e. such that

$$\sup_{s\leq t} \int_X (1 + E^\beta(y))\mu_s(dy) < \infty \tag{1.7}$$

and

$$\int_X E(y)\mu_t(dy) = \int_X E(y)\mu_0(dy) \tag{1.8}$$

for all $t \geq 0$.

The first objective of this paper is to establish the corresponding central limit theorem (CLT), i.e. to show that the process

$$F_t^h(Z_0^h, \mu_0) = h^{-1/2}(Z_t^h(Z_0^h) - \mu_t(\mu_0))$$

of normalized fluctuations of Z_t^h around its dynamic law of large numbers μ_t converges in some sense to a generalized Gaussian Ornstein-Uhlenbeck process on $\mathcal{M}(X)$ or a more general space of distributions. We obtain this result under some mild technical assumptions on the coagulation kernel thus presenting a solution to the problem 10 from the list of open problems on coagulation from [1].

It is worth noting that though for the classical processes preserving the number of particles (like interacting diffusions or Boltzmann type collisions) the results of CLT type are well established and widely presented in the literature (see e.g. [13], [6], [29]) and references therein), for the processes with a random number of particles the work on CLT began recently. For coagulation processes with discrete state space $X = \mathbf{N}$ and uniformly bounded intensities the central limit for fluctuations was obtained in [8] using stochastic calculus. For general processes of coagulation, fragmentation and collisions on $X = \mathbf{R}_+$, but again with bounded intensities, the central limit was proved by a different method in [21], namely by analytic methods of the theory of semigroups. The results of the present paper are obtained by developing further the approach from [21].

The second objective of the paper is to provide precise estimates of the error term both in LLN and CLT for a wide class of bounded and unbounded functionals on measures. Note that the usual "prove compactness in the Skorohod space and choose a converging subsequence" probabilistic method does not provide such estimates (see, however, [13] for a progress in this direction for interacting diffusions).

The method of [29] (that goes back to [32] and [30]) is based on the direct study of the solutions to the infinite-dimensional Langevin equation, describing the fluctuation process, in the strong sense. A courageous attempt to work with the Smoluchovski equation (and thus with processes changing the number of particles) on the same basis was made in [8], where extreme technical difficulties were met forcing the authors to reduce their analysis not only to bounded coefficients but even to only discrete mass distributions. Our approach enhances probabilistic tools by the sound analytic input yielding the convergence of semigroups as an intermediate step before embarking on the probabilistic analysis of the distribution of fluctuations in the appropriate Hilbert space extensions of the space of Borel measures. The main novelty of our approach (both technically and methodologically) lies in the systematic study of the derivatives of the solutions to kinetic equations with respect to initial data (this approach is inspired by the analysis of such derivatives for the Boltzmann equation in [19]). The existence and regularity of these derivatives in weighted spaces of functions and measures are analyzed and the validity of CLT is proved to be connected with a certain kind of stability of these derivatives. The estimates obtained are rather subtle, the main technical stuff being developed in Sections 5 and 7. Ideologically the corner stones of our analysis are positivity, duality and perturbation techniques. On the technical side, we are led to the heavy use of appropriately chosen Banach scales, in particular weighted Sobolev spaces, which one would expect for this kind of problems (see e.g. various Banach scales in [28] for the analysis of the Schrödinger equation, [2] for quantum stochastic setting and [30], [29] for classical interacting particles). The key points of our analysis are well reflected in the step-by-step breakdown of our results in the sequence of theorems formulated in the next section, showing the steady progress in the strengthening of the convergence and estimates via (i) error term estimates for the LLN, (ii) convergence of the linear functionals of the fluctuating measures with precise rates of convergence, (iii) the semigroup convergence, (iv) the convergence of finite-dimensional approximations and finally (v) the convergence of the distributions on trajectories, each step having its peculiarity and specific technical issues in this infinite-dimensional setting as opposed to more or less straightforward connection of similar kind of results in the case of usual Feller semigroups.

The final estimates and their proofs depend on the structure and the regularity properties of the coagulation kernel. We demonstrate various aspects of our approach analyzing the following three classes of kernels:

(C1) $K(x_1, x_2) = C(E(x_1) + E(x_2))$.

Remark. This is a warming up example, for the solutions to the main equations are given more or less explicitly in this case (see Proof of Proposition 5.2).

(C2) $K(x_1, x_2) \leq C(1 + \sqrt{E(x_1)})(1 + \sqrt{E(x_2)})$.

Remark. This model is analyzed to show the kind of results one can expect to obtain without assuming any differential or linear structure on the state space X. The unavoidable shortcoming of these results is connected with the absence of an appropriate space of generalized functions to work with. Hence the estimate of errors in LLN and CLT have to depend on something like the norm of $F_0^h = (Z_0^h - \mu_0)/\sqrt{h}$ in $\mathcal{M}(X)$. But general μ_0 can not be approximated by Dirac measures Z_0^h in such a way that F_0^h be bounded in $\mathcal{M}(X)$. Hence the possibility to apply these results beyond discrete supported initial measures μ_0 is rather reduced, so that there is no big loss in assuming (C2), which is a stronger constraint than (1.1). On the other hand, these results are open to extensions to very general spaces.

(C3) $X = \mathbf{R}_+$, $K(x_1, x_2, dy) = K(x_1, x_2)\delta(y - x_1 - x_2)$, $E(x) = x$, K is non-decreasing in each argument 2-times continuously differentiable on $(\mathbf{R}_+)^2$ up to the boundary with all the first and second partial derivatives being bounded by a constant C.

Remark. This is the case of our main interest. Unlike previous cases the estimate here turns out to depend on the norm of F_0^h coming from the dual space to continuously differentiable functions, and this norm can be easily made small for an arbitrary measure μ_0 on X. Therefore, to shorten the exposition, we shall prove CLT completely, up to the convergence of the distributions of processes on the Skorohod space of càdlàg functions, only for this case, restricting the discussion of the first two cases only to the convergence of linear functionals. For simplicity, we choose here the state space $X = \mathbf{R}_+$ of the standard Smoluchowski model, the extensions to finite-dimensional Euclidean spaces X being not difficult to obtain. Similarly we choose very strong assumptions on the derivatives (in particular, the kernels $K(x, y) = x^\alpha + y^\alpha$ with $\alpha \in (0, 1)$ are excluded by our assumption, as the derivatives of this K have a singularity at the origin). Finally let us stress that all kernels from (C1)-(C3) clearly satisfy (1.1) (possibly up to a constant multiplier).

We refer to reviews [1] and [24] for a general background in coagulation models, and to [15] for simulation and numerical methods.

The content of the paper is the following. In the next section we formulate the main results, and other sections are devoted to their proofs. In particular, Sections 4 and 5 are devoted to a detailed analysis of the equation in variations (linear approximation) around the solution of kinetic equation (1.5) that describes the derivatives of the solution to (1.5) with respect to the initial measure μ_0. At the end of Sect. 5 a new property of the kinetic equation itself is established that is crucial to our proof of CLT, but seems to be also of independent interest. Namely Propositions 5.5, 5.8 show that the solution depends Lipschitz continuously on the initial measure in the topology of the dual to the weighted spaces of continuously differentiable functions or certain weighted Sobolev spaces. In Appendix A we describe our key notations for weighted spaces of functions and distributions. In Appendix B three general result are presented (on variational derivatives, on the linear transformation of Feller processes and on the dynamics of total variations of measures), used in our proofs and placed separately in order not to interrupt the main line of arguments. In Appendix C some auxiliary facts on the evolutions specified by unbounded integral generators are presented. Though they should be essentially known to probabilists dealing with jump processes, the author did not find an appropriate reference.

To conclude the introduction we shall fix other important notations concerning variational derivatives and propagators.

For a function F on $\mathcal{M}_f(X)$ the variational derivative δF is defined by

$$\delta F(Y;x) = \lim_{s\to 0_+} \frac{1}{s}(F(Y+s\delta_x) - F(Y)),$$

where $\lim_{s\to 0_+}$ means the limit over positive s. Occasionally we shall omit the last argument here writing $\delta F(Y)$ instead of $\delta F(Y;.)$. The higher derivatives $\delta^l F(Y; x_1, ..., x_l)$ are defined inductively.

As it follows from the definition, if $\delta F(Y;.)$ exists and depends continuously on Y in the weak topology of $\mathcal{M}(X)$ (or any $\mathcal{M}_f(X)$), then the function $F(Y+s\delta_x)$ of $s \in \mathbf{R}_+$ has a continuous right derivative everywhere and hence is continuously differentiable, which implies that

$$F(Y+\delta_x) - F(Y) = \int_0^1 \delta F(Y+s\delta_x; x)\, ds. \tag{1.9}$$

We shall need an extension of this identity for more general measures in the place of the Dirac measure δ_x. To this end the following definitions turn out to be useful. For two non-negative continuous functions ϕ, f, $f(x) \to \infty$ as $x \to \infty$, we say that F belongs to $C^l(\mathcal{M}_f(X), \phi)$, $l = 0, 2, \dots,$ if $F \in C(\mathcal{M}_f(X))$ and for all $k = 1, ..., l$, $\delta^k F(Y; x_1, \dots, x_k)$ exists for all $x_1, \dots, x_k \in X^k$, $Y \in \mathcal{M}_f(X)$ and represents a continuous mapping $\mathcal{M}_f(X) \mapsto C^{sym}_{\phi\otimes\cdots\otimes\phi,\infty}(X^k)$, where $\mathcal{M}_f(X)$ is considered in its weak topology. We shall write shortly $C^l(\mathcal{M}_f(X))$ for $C^l(\mathcal{M}_f(X), f)$. All necessary formulae on the variational derivatives in these classes are collected in Lemma B.1.

We shall sometimes omit X from the notation writing shortly, say, $\mathcal{M}_f$ instead of $\mathcal{M}_f(X)$, which should not lead to ambiguity.

Remark. The introduction of the cumbersome notations $C^m(\mathcal{M}_f(X), \phi)$ is motivated by the fact that (under our assumption on the growth of the coagulation rates) if one considers the solution to the kinetic equations μ_t with $\mu_0 \in \mathcal{M}_{1+E^\beta}$, then usually $\dot{\mu}_t \in \mathcal{M}_{1+E^{\beta-1}}$ and the derivatives of μ_t with respect to the initial data belong to C_{1+E^k} with certain $k < \beta$, see Sections 4 and 5.

If S_t is a family of topological linear spaces, $t \in \mathbf{R}^+$, we shall say that a family of continuous linear operators $U^{t,r} : S^r \mapsto S^t$, $r \le t$ (respectively $t \le r$) is a propagator (respectively a backward propagator), if $U^{t,t}$ is the identity operator in S^t for all t and the following propagator equation (called Chapman-Kolmogorov equation in the probabilistic context) holds for $r \le s \le t$ (respectively for $t \le s \le r$):

$$U^{t,s}U^{s,r} = U^{t,r}. \tag{1.10}$$

By c and κ we shall denote various constants indicating in brackets (when appropriate) the parameters on which they depend.

For an operator U in a Banach space B we shall denote by $\|U\|_B$ the norm of U as a bounded linear operator in B.

At last, we shall use occasionally the obvious formula

$$\sum_{I\subset\{1,\dots,n\},|I|=2} f(\mathbf{x}_I) = \frac{1}{2}\int\int f(z_1,z_2)\delta_{\mathbf{x}}(dz_1)\delta_{\mathbf{x}}(dz_2) - \frac{1}{2}\int f(z,z)\delta_{\mathbf{x}}(dz), \tag{1.11}$$

valid for any $f \in C^{sym}(X^2)$ and $\mathbf{x} = (x_1, \ldots, x_n) \in X^n$.

2 Results

In the formulations of our main results below (as well as throughout the text) we freely use the notations for weighted norms and spaces described in detail in Appendix A.

First we recall some known results on the Cauchy problem for equation (1.5). A proof of the following two results can be found in [34] and [18] respectively. Recall that we always assume that our continuous coagulation kernel $K(x_1, x_2; dy)$ preserves E and enjoys the estimate (1.1).

Proposition 2.1 *If a finite measure μ_0 has a finite moment of order $\beta \geq 2$, i.e. if*

$$\int_X (1 + E^\beta(y))\mu_0(dy) < \infty, \tag{2.1}$$

then equation (1.5) *has a unique solution μ_t with the initial condition μ_0 satisfying* (1.7) *and* (1.8) *for arbitrary t. Moreover,*

$$\sup_{s\leq t} \int_X E^\beta(y)\mu_s(dy) \leq c(C, t, \beta, (1 + E, \mu_0))(E^\beta, \mu_0) \tag{2.2}$$

with a constant c, and the mapping $\mu_0 \mapsto \mu_t$ is Lipschitz continuous so that

$$\sup_{s\leq t} \|\mu_s(\mu_0^1) - \mu_s(\mu_0^2)\|_{1+E^\omega} \leq c(C, t, \beta, (1+E, \mu_0^1+\mu_0^2))(1+E^{1+\omega}, \mu_0^1+\mu_0^2)\|\mu_0^1 - \mu_0^2\|_{1+E^\omega} \tag{2.3}$$

for any $\omega \in [1, \beta - 1]$.

Proposition 2.2 *Solutions μ_t from the previous Proposition enjoy the following regularity properties:*

(i) for any $g \in B_{1+E^\beta,\infty}$ (respectively $g \in B_{1+E^{\beta-1},\infty}$) the function $\int g(x)\mu_t(dx)$ is a continuous function of t (respectively continuously differentiable function of t and (1.5) *holds);*

(ii) the function $t \mapsto \mu_t$ is absolutely continuous in the norm topology of $\mathcal{M}_{1+E^{\beta-1}}(X)$ and is continuously differentiable and satisfies the strong version of (1.5) *in the norm topology of $\mathcal{M}_{1+E^{\beta-\gamma}}(X)$ for any $\gamma \in (1, \beta]$.*

Remarks.

1. The basic ideas of proving Proposition 2.1 go back to the analysis of the Boltzmann equation in [35]. Formulas (2.2), (2.3) are proved in [34] only for $\beta = 2$ and $\omega = 1$ respectively, but the above extension is straightforward.

2. Statement (ii) of Proposition 2.2 is proved in [18] only for $\gamma = \beta$, but the extension given above is straightforward. In fact (ii) is done in the same way as the similar statement of Theorem C.2 from Appendix.

It is worth to observe that the operator L_h has the form of the r.h.s. of equation (C.1) from the Appendix with $\mathcal{M}_{h\delta}$ instead of X and with the (time homogeneous) intensity

$$a(h\delta_{\mathbf{x}}) = h \sum_{I\subset\{1,\ldots,n\}:|I|=2} \int K(\mathbf{x}_I; dy) \leq 3Ch^{-1}(1+E, h\delta_{\mathbf{x}})(1, h\delta_{\mathbf{x}}). \tag{2.4}$$

As the jumps in (1.4) increase neither $(1, h\delta_{\mathbf{x}}))$ nor $(E, h\delta_{\mathbf{x}})$, it is convenient to consider the process Z_h^t on a reduced state space

$$\mathcal{M}_{h\delta}^{e_0,e_1} = \{Y \in \mathcal{M}_{h\delta} : (1, Y) \leq e_0, (E, Y) \leq e_1\}.$$

On this reduced space the intensity (2.4) is bounded (not uniformly in h). Hence L_h is bounded in $C(\mathcal{M}_{h\delta}^{e_0,e_1})$ and generates a strongly continuous semigroup of contractions there, which we shall denote by T_t^h.

Let T_t be a semigroup specified by the solution of (1.5), i.e. $T_t f(\mu) = f(\mu_t)$, where μ_t is the solution of (1.5) with the initial condition μ given by Proposition 2.1 with some $\beta \geq 2$. We can formulate now our first result. Recall again that all notations for weighted norms used below are given in Appendix A.

Theorem 2.1 *[The rate of convergence in LLN] Let g be a continuous symmetric function on X^m and $F(Y) = (g, Y^{\otimes m})$. Assume $Y = h\delta_{\mathbf{x}}$ belongs to $\mathcal{M}_{h\delta}^{e_0,e_1}$, where $\mathbf{x} = (x_1, \ldots, x_n)$. Then under the condition (C1) or (C2)*

$$\begin{aligned}&\sup_{s\leq t}|T_t^h F(Y) - T_t F(Y)| \\ &\qquad \leq h\kappa(C, m, k, t, e_0, e_1)\|g\|_{(1+E^k)^{\otimes m}}(1 + E^{2k+2}, Y)(1 + E^k, Y)^{m-1}\end{aligned} \tag{2.5}$$

for any $k \geq 1$ and under the condition (C3)

$$\begin{aligned}&\sup_{s\leq t}|T_t^h F(Y) - T_t F(Y)| \\ &\qquad \leq h\kappa(C, m, k, t, e_0, e_1)\|g\|_{C^{2,sym}_{(1+E^k)^{\otimes m}}(X^m)}(1 + E^{2k+3}, Y)(1 + E^k, Y)^{m-1}\end{aligned} \tag{2.6}$$

for any $k \geq 0$ with a constant κ.

Remarks.

1. We give the hierarchy of estimates for the error term making precise an intuitively clear fact that the power of growth of the polynomial functions on measures for which LLN can be established depends on the order of the finite moments of the initial measure. In Section 7 we prove the same estimates (2.5), (2.6) for more general functionals F (not necessarily polynomial).

2. The estimates in case (C2) can be improved. However, not going into this detail allows one to keep unified formulae for cases (C1) and (C2).

The idea of the proof of this theorem is based on the representation

$$T_tF(Y) - T_t^hF(Y) = \int_0^t T_{t-s}^h(L_h - \mathcal{L})T_sF(Y)\,ds \tag{2.7}$$

for the l.h.s. of (2.5), (2.6), where $\mathcal{L}$ is the generator of the deterministic semigroup $T_t(Y) = Y_t$ that yields the solution to the Cauchy problem of Smoluchovski kinetic equation (1.5). It turns out further (see Section 3) that this difference is expressed in terms of the variational derivatives of Y_t with respect to the initial data Y. Analysis of those derivatives is then carried out via the solutions to the system in variations (or linearization) of Smoluchovski equation around its solution. This rather heavy analysis with variety of necessary estimates in different norms is carried out in Sections 4 and 5, Theorem 2.1 being finally proved in Section 6.

Recall that

$$F_t^h(Z_0^h, \mu_0) = h^{-1/2}(Z_t^h(Z_0^h) - \mu_t(\mu_0))$$

is the process of the normalized fluctuations. The main goal of this paper is to prove that as $h \to 0$ this process converges to the generalized Gaussian Ornstein-Uhlenbeck (OU) measure-valued process with the (non-homogeneous) generator

$$\begin{aligned}\Lambda_tF(Y) = \frac{1}{2}\int\int\int(\delta F(Y), \delta_y - \delta_{z_1} - \delta_{z_2})K(z_1, z_2; dy)(Y(dz_1)\mu_t(dz_2) + \mu_t(dz_1)Y(dz_2))\\ + \frac{1}{4}\int\int\int(\delta^2F(Y), (\delta_y - \delta_{z_1} - \delta_{z_2})^{\otimes 2})K(z_1, z_2; dy)\mu_t(dz_1)\mu_t(dz_2).\end{aligned} \tag{2.8}$$

The formal calculation of this generator is not difficult and is carried out in Section 3. A rigorous construction of the corresponding OU process already requires some work. Namely, the corresponding semigroup will be obtained in Theorem 2.3 on the set of cylinder functions with the existence of the process itself following then from the tightness of the approximations (obtained in Theorem 2.6) together with the uniquely specified finite dimensional limiting distributions.

The generalized infinite dimensional Ornstein- Uhlenbeck processes and the corresponding Mehler semigroups represent a widely discussed topic in the current mathematical literature, see e.g. [23] and references therein for general theory, [33] for some properties of Gaussian Mehler semigroups and [7] for the connection with branching processes with immigration. The peculiarity of the process we are dealing with lies in its 'growing coefficients'. We shall analyze this process by the analytic tools developed in Sections 4 and 5. Let us start its discussion with an obvious observation that the polynomial functionals of the form $F(Y) = (g, Y^{\otimes m})$, $g \in C^{sym}(X^m)$, on measures are invariant under Λ_t. In particular, for a linear functional $F(Y) = (g, Y)$ (i.e. for $m = 1$)

$$\Lambda_tF(Y) = \frac{1}{2}\int\int\int(g(y) - g(z_1) - g(z_2))K(z_1, z_2; dy)(Y(dz_1)\mu_t(dz_2) + \mu_t(dz_1)Y(dz_2)). \tag{2.9}$$

Hence the evolution (in the inverse time) of the linear functionals specified by the equation $\dot{F}_t = -\Lambda_tF_t$, $F_t(Y) = (g_t, Y)$ can be described by the equation

$$\dot{g}(z) = -\Lambda_tg(z) = -\int\int(g(y) - g(x) - g(z))K(x, z; dy)\mu_t(dx) \tag{2.10}$$

on the coefficient functions g_t (with some abuse of notation we denoted the action of Λ_t on the coefficient functions again by Λ_t). Let $U^{t,r}$ be the backward propagator of this equation, i.e. the resolving operator of the Cauchy problem $\dot{g} = -\Lambda_t g$ for $t \leq r$ with a given g_r. As we shall show in Propositions 5.2 - 5.4, the evolution $U^{t,r}$ is well defined in $C_{1+E^k}(X)$ in cases (C1)-(C2), and in $C^{2,0}_{1+E^k}(X)$ in case (C3).

We shall formulate now various versions of CLT that are of interest in their own right (in particular, due to the precise rates of the error terms), but also reflect the steps of proving the last (and more advanced) version on the convergence of the distributions on the trajectories of the fluctuation process.

The next result describes the (trace of) CLT on linear functions. Though this is a sort of reduced CLT, as it 'does not feel' the quadratic part of the generator of the limiting Gaussian process, technically it is the major ingredient for proving further advanced versions: convergence of semigroups, convergence of finite dimensional distributions, convergence of the distributions of trajectories in the weak sense, and the final formulation of Theorem 2.6.

Theorem 2.2 *[reduced CLT: convergence of linear functionals] Under condition (C1) or (C2)*

$$\sup_{s\leq t} |\mathbf{E}(g, F^h_s(Z^h_0,\mu_0)) - (U^{0,s}g, F^h_0(Z^h_0,\mu_0))|$$
$$\leq \kappa(C,t,k,e_0,e_1)\sqrt{h}\|g\|_{1+E^k}(1+E^{2k+2}, Z^h_0+\mu_0)^2 \left(1 + \left\|F^h_0(Z^h_0,\mu_0)\right\|^2_{\mathcal{M}_{1+E^{k+1}}(X)}\right) \quad (2.11)$$

for all $k \geq 1$, $g \in C_{1+E^k}(X)$, and under condition (C3)

$$\sup_{s\leq t} |\mathbf{E}(g, F^h_s(Z^h_0,\mu_0)) - (U^{0,s}g, F^h_0)|$$
$$\leq \kappa(C,t,k,e_0,e_1)\sqrt{h}\|g\|_{C^{2,0}_{1+E^k}}(1+E^{2k+3}, Z^h_0+\mu_0)^2 \left(1 + \left\|F^h_0(Z^h_0,\mu_0)\right\|^2_{\mathcal{M}^1_{1+E^{k+1}}(X)}\right) \quad (2.12)$$

for all $k \geq 0$, $g \in C^{2,0}_{1+E^k}(X)$, where the bald $\mathbf{E}$ denotes the expectation with respect to the process Z^h_t.

A proof of this theorem is given in Section 6. It is a consequence of Theorem 2.1 and the estimate on the second moment of the fluctuation process obtained in Proposition 7.1.

To shorten the exposition, the further more refined versions of CLT will be given only in the most important case (C3). Though all the results have natural modifications in cases (C1) and (C2), let us stress again that for their applicability in cases (C1), (C2) one needs the initial fluctuation F^h_0 to be bounded in the norm of $\mathcal{M}_{1+E^{k+1}}(X)$, which is possible basically only for discrete initial distributions μ_0.

For our purposes it will be enough to construct the propagator of the equation $\dot{F} = -\Lambda_t F$ only on the set of cylinder functions $\mathcal{C}^n_k = \mathcal{C}^n_k(\mathcal{M}^m_{1+E^k})$, $m = 1, 2$, on measures that have the form

$$\Phi^{\phi_1,...,\phi_n}_f(Y) = f((\phi_1, Y), ..., (\phi_n, Y)) \quad (2.13)$$

with $f \in C(\mathbf{R}^n)$, and $\phi_1, ..., \phi_n \in C^{m,0}_{1+E^k}$. By $\mathcal{C}_k$ we shall denote the union of $\mathcal{C}^n_k$ for all $n = 0, 1, ...$ (of course, functions from $\mathcal{C}^0_k$ are just constants). Similarly one defines the cylinder functions $\mathcal{C}^n_k((L^{m,0}_{2,1+E^k})')$ under condition (C3).

The Banach space of k times continuously differentiable functions on $\mathbf{R}^d$ (with the norm being the maximum of the sup-norms of a function and all its partial derivative up to and including the order k) will be denoted, as usual, by $C^k(\mathbf{R}^d)$.

Theorem 2.3 *[limiting Mehler propagator] Under the condition (C3) for any $k \geq 0$ and a μ_0 such that $(1+E^{k+1}, \mu_0) < \infty$ there exists a propagator $OU^{t,r}$ of contractions on $\mathcal{C}_k$ preserving the subspaces $\mathcal{C}_k^n$, $n = 0,1,2,...$ such that $OU^{t,r}F$, $F \in \mathcal{C}_k$, depends continuously on t in the topology of the uniform convergence on bounded subsets of $\mathcal{M}^m_{1+E^k}$, $m = 1,2$ (respectively also of $(L^{m,0}_{2,1+E^k})'$ in case $k > 1/2$) and solves the equation $\dot{F} = -\Lambda_t F$ in the sense that if $f \in C^2(\mathbf{R}^d)$ in* (2.13), *then*

$$\frac{d}{dt} OU^{t,r} \Phi_f^{\phi_1,..,\phi_n}(Y) = -\Lambda_t OU^{t,r} \Phi_f^{\phi_1,..,\phi_n}(Y), \quad 0 \leq t \leq r, \tag{2.14}$$

uniformly for Y from bounded subsets of $\mathcal{M}^m_{1+E^k}$ (respectively $(L^{m,0}_{2,1+E^k})'$).

Our goal is to prove that this generalized infinite-dimensional Ornstein-Uhlenbeck (or Mehler) semigroup describes the limiting Gaussian distributions of the fluctuation process F_t^h.

Theorem 2.4 *[CLT: convergence of semigroups] Suppose $k \geq 0$ and $h_0 > 0$ are given such that*

$$\sup_{h \leq h_0} (1 + E^{2k+5}, Z_0^h + \mu_0) < \infty. \tag{2.15}$$

(i) Let $\Phi \in \mathcal{C}_k^n(\mathcal{M}^2_{1+E^k})$ be given by (2.13) *with $f \in C^3(\mathbf{R}^n)$ and all $\phi_j \in C^{2,0}_{1+E^k}(X)$. Then*

$$\sup_{s \leq t} |\mathbf{E}\Phi(F_t^h(Z_0^h, \mu_0)) - OU^{0,t}\Phi(F_0^h)|$$

$$\leq \kappa(C,t,k,e_0,e_1)\sqrt{h} \max_j \|\phi_j\|_{C^{2,0}_{1+E^k}} \|f\|_{C^3(\mathbf{R}^n)} (1+E^{2k+5}, Z_0^h+\mu_0)^2 \left(1 + \|F_0^h\|^2_{\mathcal{M}^1_{1+E^{k+1}}(X)}\right). \tag{2.16}$$

(ii) If $\Phi \in \mathcal{C}_k^n(\mathcal{M}^2_{1+E^k})$ (with not necessarily smooth f in the representation (2.13)*) and F_0^h converges to some F_0 as $h \to 0$ in the $\star$-weak topology of $\mathcal{M}^1_{1+E^{k+1}}$, then*

$$\lim_{h \to 0} |\mathbf{E}\Phi(F_t^h(Z_0^h, \mu_0)) - OU^{0,t}\Phi(F_0)| = 0 \tag{2.17}$$

uniformly for F_0^h from a bounded subset of $\mathcal{M}^1_{1+E^{k+1}}$ and t from a compact interval.

Theorem 2.5 *[CLT: convergence of finite dimensional distributions] Suppose* (2.15) *holds, $\phi_1,...,\phi_n \in C^{2,0}_{1+E^k}(\mathbf{R}_+)$ and $F_0^h \in (L^{2,0}_{2,1+E^{k+2}})'$ converges to some F_0 in $(L^{2,0}_{2,1+E^{k+2}})'$, as $h \to 0$. Then the $\mathbf{R}^n$-valued random variables*

$$\Phi^h_{t_1,...,t_n} = ((\phi_1, F^h_{t_1}(Z_0^h, \mu_0)), ..., (\phi_n, F^h_{t_n}(Z_0^h, \mu_0))), \quad 0 < t_1 \leq ... \leq t_n,$$

converge in distribution, as $h \to 0$, to a Gaussian random variable with the characteristic function

$$g_{t_1,...,t_n}(p_1,...,p_n) = \exp\{i \sum_{j=1}^n p_j (U^{0,t_j}\phi_j, F_0) - \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \sum_{l,k=j}^n p_l p_k \Pi(s, U^{s,t_l}\phi_l, U^{s,t_k}\phi_k)\, ds\}, \tag{2.18}$$

where $t_0 = 0$ *and*

$$\Pi(t,\phi,\psi) = \frac{1}{4}\int\int\int(\phi\otimes\psi,(\delta_y-\delta_{z_1}-\delta_{z_2})^{\otimes 2})K(z_1,z_2;dy)\mu_t(dz_1)\mu_t(dz_2). \qquad (2.19)$$

In particular, for $t = t_1 = ... = t_n$ *it implies*

$$\lim_{h\to 0}\mathbf{E}\exp\{i\sum_{j=1}^n(\phi_j,F_t^h)\} = \exp\{i\sum_{j=1}^n(U^{0,t}\phi_j,F_0) - \sum_{j,k=1}^n\int_0^t\Pi(s,U^{s,t}\phi_j,U^{s,t}\phi_k)\,ds\}.$$

Note that passing from Theorem 2.4 to Theorem 2.5 would be automatic for finite dimensional Feller processes, but in our infinite dimensional setting this is not at all straightforward and requires additional use of the Hilbert space methods leading to some uniform bounds on the process of fluctuation obtained in Section 7.

Theorem 2.6 *[CLT: convergence of the process of fluctuations] Suppose the conditions of Theorems 2.4, 2.5 hold. (i) For any* $\phi \in C^{2,0}_{1+E^k}(\mathbf{R}_+)$ *the real valued processes* $(\phi, F_t^h(Z_0^h,\mu_0))$ *converge in the sense of the distribution in the Skorohod space of càdlàg functions (equipped with its standard* J_1*-topology) to the Gaussian process with finite-dimensional distributions specified by Theorem 2.5. (ii) The process of fluctuations* $F_t^h(Z_0^h,\mu_0)$ *converges in distributions on the Skorohod space of càdlàg functions* $D([0,T];(L^{2,0}_{2,1+E^{k+2}}(\mathbf{R}_+))')$ *(with* J_1*-topology), where* $(L^{2,0}_{2,1+E^{k+2}}(\mathbf{R}_+))'$ *is considered in its weak topology, to a Gaussian process with finite-dimensional distributions specified by Theorem 2.5.*

Proof is given at the end of Section 8. Let us note only that once the previous results are obtained everything that remains to prove for Theorem 2.6 is the tightness of the approximations, i.e. the existence of a limiting point, because all finite dimensional distributions of such a point are already uniquely specified by Theorems 2.4, 2.5.

3 Calculations of generators

From now on we denote by $\mu_t = \mu_t(\mu_0)$ the solution to (1.5) given by Proposition 2.1 with a $\beta \geq 2$. To begin with, let us extend the action of T_t^h beyond the space $C(\mathcal{M}_{h\delta}^{e_0,e_1})$.

Proposition 3.1 *For any positive* e_0, e_1 *and* $1 \leq l \leq m$ *the operator* L_h *is bounded in the space* $C_{(1+E^l,\cdot)^m}(\mathcal{M}_{h\delta}^{e_0,e_1})$ *and defines a strongly continuous semigroup there (again denoted by* T_t^h*) such that*

$$\|T_t^h\|_{C_{(1+E^l,\cdot)^m}(\mathcal{M}_{h\delta}^{e_0,e_1})} \leq \exp\{c(C,m,l)e_1 t\}. \qquad (3.1)$$

Proof. Let us show that

$$L_h F(Y) \leq c(C,m,l)e_1 F(Y) \qquad (3.2)$$

for $Y = h\delta_{\mathbf{x}}$ and $F(Y) = (1+E^l, Y)^m$. One has

$$L_h F(Y) = h\sum_{I\subset\{1,...,n\}:|I|=2}\int[(1+E^l,Y+h(\delta_y-\delta_{\mathbf{x}_I}))^m-(1+E^l,Y)^m]K(\mathbf{x}_I;dy).$$

As

$$(1+E^l, h(\delta_y-\delta_{x_i}-\delta_{x_j})) \le h[(E(x_i)+E(x_j))^l - E^l(x_i)-E^l(x_j)]$$
$$\le hc(l)[E^{l-1}(x_i)E(x_j)+E(x_i)E^{l-1}(x_j)]$$

and using the obvious inequality $(a+b)^m - a^m \le c(m)(a^{m-1}b+b^m)$ one obtains

$$\begin{aligned}
L_hF(Y) \le hc(m,l) &\sum_{I\subset\{1,\dots,n\}:|I|=2} [(1+E^l,Y)^{m-1}h(E^{l-1}(x_i)E(x_j)+E(x_i)E^{l-1}(x_j)) \\
&+ h^m(E^{l-1}(x_i)E(x_j)+E(x_i)E^{l-1}(x_j))^m]K(\mathbf{x}_I;dy) \\
&\le c(C,m,l)\int\int[(1+E^l,Y)^{m-1}(E^{l-1}(z_1)E(z_2)+E(z_1)E^{l-1}(z_2)) \\
&+ h^{m-1}(E^{l-1}(z_1)E(z_2)+E(z_1)E^{l-1}(z_2))^m](1+E(z_1)+E(z_2))Y(dz_1)Y(dz_2),
\end{aligned}$$

where we used (1.11). By symmetry it is enough to estimate the integral over the set where $E(z_1)\ge E(z_2)$. Consequently $L_hF(Y)$ does not exceed

$$\begin{aligned}
c\int&[(1+E^l,Y)^{m-1}E^{l-1}(z_1)E(z_2)+h^{m-1}(E^{l-1}(z_1)E(z_2))^m](1+E(z_1))Y(dz_1)Y(dz_2) \\
&\le c(1+E^l,Y)^m(E,Y)+h^{m-1}c\int E^{m(l-1)+1}(z_1)E^m(z_2)Y(dz_1)Y(dz_2).
\end{aligned}$$

To prove (3.2) it remains to show that the second term in the last expression can be estimated by its first term. This follows from the estimates:

$$\begin{aligned}
&(E^m,Y) = h\sum E^m(x_i) \le h\left(\sum E^l(x_i)\right)^{m/l} = h^{1-m/l}(E^l,Y)^{m/l}, \\
&(E^{m(l-1)+1},Y) \le h^{-1}(E^{m(l-1)},Y)(E,Y) \le h^{-m(1-1/l)}(E^l,Y)^{m(1-1/l)}(E,Y).
\end{aligned}$$

Once (3.2) is proved it follows from (2.4) that L_h is bounded in $C_{(1+E^l,\cdot)^m}(\mathcal{M}_{h\delta}^{e_0,e_1})$, and (3.1) follows from Theorem C.1 (or Proposition C.1).

The following statement is a straightforward extension of the previous one.

Proposition 3.2 *The statement of Proposition 3.1 remains true if instead of the space $C_{(1+E^l,\cdot)^m}$ one takes a more general space $C_{(1+E^{l_1},\cdot)^{m_1}\cdots(1+E^{l_j},\cdot)^{m_j}}$.*

Next we shall calculate the generator $\mathcal{L}$ of the deterministic semigroup T_t and compare it with L_h.

Proposition 3.3 *(i) If $F\in C^1(\mathcal{M}_{1+E^\beta}(X), 1+E^{\beta-1})$, then*

$$\frac{d}{dt}T_tF(\mu_0) = \frac{d}{dt}F(\mu_t) = \mathcal{L}F(\mu_t), \tag{3.3}$$

with

$$\mathcal{L}F(Y) = \frac{1}{2}\int_X\int_{X\times X}(\delta F(Y;y)-\delta F(Y;x_1)-\delta F(Y;x_2))K(x_1,x_2;dy)Y(dx_1)Y(dx_2). \tag{3.4}$$

(ii) If the variational derivative $\delta^2 F(Y;x,y)$ exists for $Y \in \mathcal{M}^+_{1+E^\beta}$ and is a continuous function of three variables (Y taken in its $\star$-weak topology), then for any $Y = h\delta_{\mathbf{x}}$

$$
\begin{aligned}
L_h F(Y) - \mathcal{L}F(Y) = -\frac{h}{2}\int\int(\delta F(Y;y) - 2\delta F(Y;z))K(z,z;dy)Y(dz) \\
+ h^3\int_0^1 (1-s)\,ds \sum_{I\subset\{1,\dots,n\}:|I|=2}\int_X (\delta^2 F(Y+sh(\delta_y-\delta_{\mathbf{x}_I});\cdot,\cdot),(\delta_y-\delta_{\mathbf{x}_I})^{\otimes 2})K(\mathbf{x}_I;dy).
\end{aligned}
\tag{3.5}
$$

.

(iii) If $F \in C(\mathcal{M}(X))$, $Y = h\delta_{\mathbf{x}}$, then

$$
\begin{aligned}
L_h F(Y) = \frac{1}{2h}\int\int\int[F(Y+h(\delta_y-\delta_{z_1}-\delta_{z_2}))-F(Y)]K(z_1,z_2;dy)Y(dz_1)Y(dz_2) \\
-\frac{1}{2}\int\int[F(Y+h(\delta_y-2\delta_z))-F(Y)]K(z,z;dy)Y(dz).
\end{aligned}
\tag{3.6}
$$

In particular, if $F(Y) = (\phi, Y)$ with a continuous function ϕ, then

$$
\begin{aligned}
L_h F(Y) = \frac{1}{2}\int\int\int[\phi(y)-\phi(z_1)-\phi(z_2)]K(z_1,z_2;dy)Y(dz_1)Y(dz_2) \\
-\frac{h}{2}\int\int[\phi(y)-2\phi(z)]K(z,z;dy)Y(dz).
\end{aligned}
\tag{3.7}
$$

Proof. (i) Follows from equation (B.3) and Proposition 2.2(i).

(ii) Applying equation (B.2)(a) to formula (1.4) yields

$$
\begin{aligned}
L_h F(Y) = h^2 \sum_{I\subset\{1,\dots,n\}:|I|=2}\int_X (\delta F(Y;\cdot),\delta_y-\delta_{\mathbf{x}_I})K(\mathbf{x}_I;dy) \\
+ h^3\int_0^1 (1-s)\,ds \sum_{I\subset\{1,\dots,n\}:|I|=2}\int_X (\delta^2 F(Y+sh(\delta_y-\delta_{\mathbf{x}_I});\cdot,\cdot),(\delta_y-\delta_{\mathbf{x}_I})^{\otimes 2})K(\mathbf{x}_I;dy).
\end{aligned}
$$

Transforming the first term of the r.h.s. of this equation by (1.11), yields (3.5).

(iii) Is obtained by applying (1.11) directly to (1.4).

Proposition 3.4 *The backward propagator*

$$U^{h;s,r}_{fl} : C(\Omega^h_r(\mathcal{M}^{e_0,e_1}_{h\delta})) \mapsto C(\Omega^h_s(\mathcal{M}^{e_0,e_1}_{h\delta}))$$

of the process of fluctuations F^h_t obtained from Z^h_t by the deterministic linear transformation $\Omega^h_t(Y) = h^{-1/2}(Y-\mu_t)$, is given by

$$U^{h;s,r}_{fl}F = (\Omega^h_s)^{-1}T^h_{r-s}\Omega^h_r F, \tag{3.8}$$

where $\Omega^h_t F(Y) = F(\Omega^h_t Y)$, and satisfies the equation

$$\frac{d}{ds}U^{h;t,s}_{fl}F = U^{h;t,s}_{fl}\Lambda^h_s F; \quad t < s < T, \tag{3.9}$$

for $F \in C^3(\mathcal{M}_{1+E^\beta,1+E^{\beta-1}}(X))$, *where*

$$
\begin{aligned}
\Lambda_t^h F(Y) &= \Lambda_t F(Y) + \frac{\sqrt{h}}{2}\int\int\int(\delta F(Y),\delta_y-\delta_{z_1}-\delta_{z_2})K(z_1,z_2;dy)Y(dz_1)Y(dz_2)\\
&\quad-\frac{\sqrt{h}}{2}\int\int(\delta F(Y),\delta_y-2\delta_z)K(z,z;dy)(\mu_t+\sqrt{h}Y)(dz)\\
&\quad+\frac{\sqrt{h}}{4}\int\int\int(\delta^2 F(Y),(\delta_y-\delta_{z_1}-\delta_{z_2})^{\otimes 2})K(z_1,z_2;dy)(Y(dz_1)\mu_t(dz_2)+Y(dz_2)\mu_t(dz_1))\\
&\quad-\frac{h}{4}\int\int\int(\delta^2 F(Y),(\delta_y-2\delta_z)^{\otimes 2})K(z,z;dy)(\mu_t+\sqrt{h}Y)(dz)\\
&\quad+\frac{h}{4}\int\int\int(\delta^2 F(Y),(\delta_y-\delta_{z_1}-\delta_{z_2})^{\otimes 2})K(z_1,z_2;dy)Y(dz_1)Y(dz_2)\\
&\quad+\frac{\sqrt{h}}{4}\int_0^1(1-s)^2\,ds\int\int\int(\delta^3 F(Y+s\sqrt{h}(\delta_y-\delta_{z_1}-\delta_{z_2}),\cdot),(\delta_y-\delta_{z_1}-\delta_{z_2})^{\otimes 3})\\
&\qquad\times K(z_1,z_2;dy)(\mu_t+\sqrt{h}Y)(dz_1)(\mu_t+\sqrt{h}Y)(dz_2)\\
&\quad-\frac{h^{3/2}}{4}\int_0^1(1-s)^2\,ds\int\int(\delta^3 F(Y+s\sqrt{h}(\delta_y-2\delta_z),\cdot),(\delta_y-2\delta_z)^{\otimes 3})K(z,z;dy)(\mu_t+\sqrt{h}Y)(dz).
\end{aligned}
\tag{3.10}
$$

Proof. According to Lemma B.2 the backward propagator $U_{fl}^{h;s,r}$ is given by (3.8) and satisfies (3.9) for $F \in C(\Omega_{[0,T]}(\mathcal{M}_{h\delta}^{e_0,e_1}))$ (see Lemma B.2 for this notation), where

$$
\Lambda_t^h\psi = (\Omega_t^h)^{-1}L_h\Omega_t^h\psi - h^{-1/2}\left(\frac{\delta\psi}{\delta Y},\dot{\mu}_t\right). \tag{3.11}
$$

Applying (3.6) yields

$$
\begin{aligned}
L_h\Omega_t^h F(Y) &= \frac{1}{2h}\int\int\int\left[F\left(\frac{Y+h(\delta_y-\delta_{z_1}-\delta_{z_2})-\mu_t}{\sqrt{h}}\right)-F\left(\frac{Y-\mu_t}{\sqrt{h}}\right)\right]\\
&\quad\times K(z_1,z_2;;dy)Y(dz_1)Y(dz_2)\\
&\quad-\frac{1}{2}\int\int\left[F\left(\frac{Y+h(\delta_y-2\delta_s)-\mu_t}{\sqrt{h}}\right)-F\left(\frac{Y-\mu_t}{\sqrt{h}}\right)\right]K(z,z;dy)Y(dz)
\end{aligned}
$$

and consequently

$$
\begin{aligned}
(\Omega_t^h)^{-1}L_h\Omega_t^h F(Y) &= \frac{1}{2h}\int\int\int\Big[(F(Y+\sqrt{h}(\delta_y-\delta_{z_1}-\delta_{z_2})-F(Y)\Big]\\
&\quad\times K(z_1,z_2;dy)(\sqrt{h}Y+\mu_t)(dz_1)(\sqrt{h}Y+\mu_t)(dz_2)\\
&\quad-\frac{1}{2}\int\int[F(Y+\sqrt{h}(\delta_y-2\delta_z))-F(Y)]K(z,z;dy)(\sqrt{h}Y+\mu_t)(dz).
\end{aligned}
\tag{3.12}
$$

Applying equation (B.2) (b) yields

$$
\begin{aligned}
F(Y+\sqrt{h}(\delta_y-\delta_{z_1}-\delta_{z_2}))-F(Y) &= \sqrt{h}(\delta F(Y),\delta_y-\delta_{z_1}-\delta_{z_2})+\frac{h}{2}(\delta^2F(Y),(\delta_y-\delta_{z_1}-\delta_{z_2})^{\otimes 2})\\
&\quad+\frac{h^{3/2}}{2}\int_0^1(1-s)^2(\delta^3F(Y+s\sqrt{h}(\delta_y-\delta_{z_1}-\delta_{z_2})),(\delta_y-\delta_{z_1}-\delta_{z_2})^{\otimes 3})\,ds.
\end{aligned}
$$

Hence developing the r.h.s. of (3.12) in h yields the term at $h^{-1/2}$ of the form

$$\frac{1}{2}\int\int\int(\delta F(Y),\delta_y-\delta_{z_1}-\delta_{z_2})K(z_1,z_2;dy)\mu_t(dz_1)\mu_t(dz_2),$$

the term at h^0 being precisely $\Lambda_t F(Y)$ given by (2.8), plus the remainder terms of order at least $h^{1/2}$. As the above term of order $h^{-1/2}$ cancels with the second term in (3.11) one obtains (3.10).

4 Derivatives with respect to initial data: existence

This section is devoted to the analysis of the derivatives of the solutions to equation (1.5) with respect to the initial data. Namely we are going to study the signed measures defined as

$$\xi_t=\xi_t(\mu_0;x;dz)=\frac{\delta\mu_t}{\delta\mu_0}(\mu_0;x;dz)=\lim_{s\to 0_+}\frac{1}{s}(\mu_t(\mu_0+s\delta_x)-\mu_t(\mu_0)). \tag{4.1}$$

We will occasionally omit some arguments in ξ_t to shorten the formulas.

Since the general known results on the derivatives of the evolution systems with respect to initial data are not applicable directly to (1.5) (due to unbounded coefficients), our strategy in this Section will be to introduce approximations with bounded kernels, apply standard results on variational derivatives to them, and then carefully pass to the limit.

To motivate the formulation of rigorous results, let us start with a short formal calculation. Differentiating formally equation (1.5) with respect to the initial measure μ_0 one obtains for ξ_t the equation

$$\frac{d}{dt}(g,\xi_t)=\int_{X\times X}\int_X(g(y)-g(x_1)-g(x_2))K(x_1,x_2;dy)\xi_t(dx_1)\mu_t(dx_2). \tag{4.2}$$

Of course, this is by no means a coincidence that this equation is dual to (2.10).

Introducing the second derivative

$$\eta_t=\eta_t(x,w)=\eta_t(\mu_0;x,w;dz)=\lim_{s\to 0_+}\frac{1}{s}(\xi_t(\mu_0+s\delta_w;x)-\xi_t(\mu_0;x)), \tag{4.3}$$

and differentiating (4.2) formally one obtains for η_t the equation

$$\begin{aligned}\frac{d}{dt}(g,\eta_t(x,w;,\cdot))=&\int_{X\times X}\int_X(g(y)-g(x_1)-g(x_2))K(x_1,x_2;dy)\\&\times[\eta_t(x,w;dx_1)\mu_t(dx_2)+\xi_t(x;dx_1)\xi_t(w;dx_2)].\end{aligned} \tag{4.4}$$

The aim of this section is to justify these calculations and to obtain rough estimates for ξ_t and η_t.

We start our analysis with a result on approximation of the solutions to kinetic equations by equations with bounded kernels. Let us introduce a cut-off kernel K_n that enjoys the same properties as K and is such that $K_n(x_1,x_2;dy)=K(x_1,x_2;dy)$ whenever $E(x_1)+E(x_2)\le n$ and $K_n(x_1,x_2)\le Cn$ everywhere.

For convenience, we shall assume $\beta>3$ everywhere in this section.

Proposition 4.1 *Let $\mu_0 \mapsto \mu_t^n$ be the solution, given by Proposition 2.1, to the equation* (1.5) *with K_n instead of K. Then $\mu_t^n \to \mu_t$ in the norm topology of $\mathcal{M}_{1+E^\omega}(X)$ with $\omega \in [1, \beta - 1)$ and $*$-weakly in $\mathcal{M}_{1+E^\beta}$ uniformly for t from compact sets.*

Proof. As the arguments given below use a rather standard trick in the theory of kinetic equations (similar ideas lead to a proof of Proposition 2.1) we shall give them only for $\omega = 1$.

Let σ_t^n denote the sign of the measure $\mu_t^n - \mu_t$ (i.e. the equivalence class of the densities of $\mu_t^n - \mu_t$ with respect to $|\mu_t^n - \mu_t|$ that equals ± 1 respectively in positive and negative parts of the Hahn decomposition of this measure) so that $|\mu_t^n - \mu_t| = \sigma_t^n(\mu_t^n - \mu_t)$. By Lemma B.3 one can choose a representative of σ_t^n (that we shall again denote by σ_t^n) in such a way that

$$(1 + E, |\mu_t^n - \mu_t|) = \int_0^t \left(\sigma_s^n(1 + E), \frac{d}{ds}(\mu_s^n - \mu_s) \right) ds. \tag{4.5}$$

Applying (1.5) one obtains from (4.5) that

$$\begin{aligned}(1 + E, |\mu_t^n - \mu_t|) = \frac{1}{2} \int_0^t ds \int [(\sigma_s^n(1 + E))(y) - (\sigma_s^n(1 + E))(x_1) - (\sigma_s^n(1 + E))(x_2)] \\ \times [K_n(x_1, x_2; dy)\mu_s^n(dx_1)\mu_s^n(dx_2) - K(x_1, x_2; dy)\mu_s(dx_1)\mu_s(dx_2)]. \end{aligned} \tag{4.6}$$

The expression in the last bracket in (4.6) can be rewritten as

$$\begin{aligned}(K_n - K)(x_1, x_2; dy)\mu_s^n(dx_1)\mu_s^n(dx_2) \\ + K(x_1, x_2; dy)[(\mu_s^n(dx_1) - \mu_s(dx_1))\mu_s^n(dx_2) + \mu_s(dx_1)(\mu_s^n(dx_2) - \mu_s(dx_2))]. \end{aligned} \tag{4.7}$$

As μ_s^n are uniformly bounded in $\mathcal{M}_{1+E^\beta}$ and

$$(1 + E(x_1) + E(x_2)) \int_X (K_n - K)(x_1, x_2; dy) \leq Cn^{-\epsilon}(1 + E(x_1) + E(x_2))^{2+\epsilon}$$

for $2 + \epsilon \leq \beta$, the contribution of the first term in (4.7) to the r.h.s. of (4.6) tends to zero as $n \to \infty$. The second and the third terms in (4.7) ar similar. Let us analyze the second term only. Its contribution to the r.h.s; of (4.6) can be written as

$$\begin{aligned}\frac{1}{2} \int_0^t ds \int [(\sigma_s^n(1 + E))(y) - (\sigma_s^n(1 + E))(x_1) - (\sigma_s^n(1 + E))(x_2)] \\ \times K(x_1, x_2; dy)\sigma_s^n(x_1)|\mu_s^n(dx_1) - \mu_s(dx_1)|\mu_s^n(dx_2), \end{aligned}$$

which does not exceed

$$\begin{aligned}\frac{1}{2} \int_0^t ds \int [(1 + E)(y) - (1 + E)(x_1) + (1 + E)(x_2)] \\ \times K(x_1, x_2; dy)|\mu_s^n(dx_1) - \mu_s(dx_1)|\mu_s^n(dx_2), \end{aligned}$$

because $(\sigma_s^n(x_1))^2 = 1$ and $|\sigma_s^n(x_j)| \leq 1$, $j = 1, 2$. Since K preserves E and (1.1) holds, the latter expression does not exceed

$$\begin{aligned}C \int_0^t ds \int (1 + E(x_2))(1 + E(x_1) + E(x_2))|\mu_s^n(dx_1) - \mu_s(dx_1)|\mu_s^n(dx_2) \\ \leq C \int_0^t ds(1 + E, |\mu_s^n - \mu_s|)\|\mu_s^n\|_{1+E^2}. \end{aligned}$$

Consequently by Gronwall's lemma one concludes that

$$\|\mu_t^n - \mu_t\|_{1+E} = (1+E, |\mu_t^n - \mu_t|) = o(1)_{n\to\infty} \exp\left\{ t \sup_{s\in[0,t]} \|\mu_s\|_{1+E^2} \right\}.$$

Finally, once the convergence in the norm topology of any $\mathcal{M}_{1+E^\gamma}$ with $\gamma > 0$ is established, the $*$-weak convergence in $\mathcal{M}_{1+E^\beta}$ follows from the uniform (in n) boundedness of μ_n.

Proposition 4.2 *(i) Under the assumptions of Proposition 2.1 the backward propagator $U^{t,r}$ of equation* (2.10) *is well defined and is strongly continuous in the space $C_{1+E^{\beta-1},\infty}(X)$. Moreover, there exists a unique solution ξ_t to* (4.2) *in the sense that $\xi_0 = \delta_x$, ξ_t is a $*$-weakly continuous function $\{t \geq 0\} \mapsto \mathcal{M}_{1+E^{\beta-1}}(X)$ and* (4.2) *holds for all $g \in C_{1+E}(X)$. Finally,*

$$\|\xi_t(\cdot; x)\|_{1+E^\omega} \leq \kappa(t, \|\mu_0\|_{1+E^{1+\omega}})(1+E^\omega)(x) \tag{4.8}$$

for all $\omega \in [1, \beta-1]$ and some constant κ, ξ_t is continuous with respect to t in the norm topology of $\mathcal{M}_{1+E^{\beta-1-\epsilon}}$ and is continuously differentiable in the norm topology of $\mathcal{M}_{1+E^{\beta-2-\epsilon}}$ for all $\epsilon > 0$.

(ii) If ξ_t^n are defined as ξ_t but from the cut-off kernels K_n with μ_t^n instead of μ_t, then $\xi_t^n \to \xi_t$, as $n \to \infty$ in the norm topology of $\mathcal{M}_{1+E^\omega}$ with $\omega \in [1, \beta-2)$ and in the $*$-weak topology of $\mathcal{M}_{1+E^{\beta-1}}$.

(iii) ξ_t depends Lipschitz continuously on μ_0 in the norm of $\mathcal{M}_{1+E^\omega}$ for $\omega \in [1, \beta-2]$ so that

$$\sup_{s\leq t} \|\xi_s(\mu_0^1) - \xi_s(\mu_0^2)\|_{1+E^\omega} \leq \kappa(C, t, e_0, e_1, (E^{2+\omega}, \mu_0^1 + \mu_0^2))\|\mu_0^1 - \mu_0^2\|_{1+E^{1+\omega}}(1 + E^{1+\omega}(x)).$$

(iv) ξ_t can be defined by the r.h.s; of (4.1) *with the limit existing in the norm topology of $\mathcal{M}_{1+E^\omega}(X)$ with $\omega \in [1, \beta-1)$ and in the $*$-weak topology of $\mathcal{M}_{1+E^{\beta-1}}$.*

Proof. (i) Equation (4.2) is dual to (2.10) and is a particular case of equation (C.14) from Appendix with

$$A_t g(x) = \int_X \int_X (g(y) - g(x)) K(z, x; dy) \mu_t(dz), \tag{4.9}$$

and

$$B_t g(x) = \int_X g(z) \int_X K(z, x; dy) \mu_t(dz). \tag{4.10}$$

In the notations of Theorem C.2 one has in our case

$$a_t(x) = \int_X \int_X K(z, x; dy) \mu_t(dz) \leq C(1 + E(x)) \|\mu_t\|_{1+E},$$

and for all $\omega \leq \beta - 1$

$$\|B_t g\|_{1+E} = \|B_t g/(1+E)\| \leq C \sup_x \{ \frac{\int g(z)(1 + E(x) + E(z))\mu_t(dz)}{1 + E(x)} \}$$

$$\leq C\|g\|_{1+E^\omega} \int (1 + E^\omega(z))(1 + E(z))\mu_t(dz) \leq 3C\|g\|_{1+E^\omega}\|\mu_t\|_{1+E^{\omega+1}}.$$

Moreover, as $\omega \geq 1$

$$A_t(1+E^\omega)(x) \leq C \int_X ((E(x)+E(z))^\omega - E^\omega(x))(1+E(z)+E(x))\mu_t(dz)$$
$$\leq Cc(\omega) \int_X (E^{\omega-1}(x)E(z)+E^\omega(z))(1+E(z)+E(x))\mu_t(dz) \leq Cc(\omega)(1+E^\omega)(x)\|\mu_t\|_{1+E^{1+\omega}}.$$

Hence the required well-posedness of the dual equations (2.10) and (4.2) and estimate (4.8) for $\omega = \beta - 1$ follow from Theorem C.2 (i), (ii) with $\psi_1 = 1+E^s$, $s \in [1, \beta-2)$, and $\psi_2 = 1+E^{\beta-1}$. The last statement of (i) follows from Theorem C.2 (iii). Estimate (4.8) for other $\omega \in [1, \beta-1]$ follows again from Theorem C.2 (i) and the estimates for a_t and B_t given above. [1]

(ii) The proof is the same as the proof of Proposition 4.1 above.

(iii) The proof of this statement is practically the same as for the corresponding statement (see Proposition 2.1(i)) for the solution of kinetic equation and uses the same trick as in the proof of Proposition 4.1 above. Namely denoting $\xi_t^j = \xi_t(\mu_0^j)$, $j = 1, 2$, one writes

$$\|\xi_t^1 - \xi_t^2\|_{1+e^\omega} = \int_0^t ds \int [(\sigma_s(1+E^\omega))(y) - (\sigma_s(1+E^\omega))(x_1) - (\sigma_s(1+E^\omega))(x_2)]$$
$$\times K(x_1, x_2; dy)[\xi_s^1(dx_1)\mu_s^1(dx_2) - \xi_s^2(dx_1)\mu_s^2(dx_2)],$$

where σ_s denotes the sign of the measure $\xi_t^1 - \xi_t^2$ (again chosen according to Lemma B.3). Next, rewriting

$$\xi_s^1(dx_1)\mu_s^1(dx_2) - \xi_s^2(dx_1)\mu_s^2(dx_2) = \sigma_s(x_1)|\xi_s^1 - \xi_s^2|(dx_1)\mu_s^1(dx_2) + \xi_s^2(dx_1)(\mu_s^1 - \mu_s^2)(dx_2)$$

one estimates from above the contribution of the first term in the above expression for $\|\xi_t^1 - \xi_t^2\|_{1+E^\omega}$ by

$$\int_0^t ds \int [E^\omega(y) - E^\omega(x_1) + E^\omega(x_2) + 1]K(x_1, x_2; dy)|\xi_s^1 - \xi_s^2|(dx_1)\mu_s^1(dx_2)$$
$$\leq c(\omega)C \int_0^t ds \int [E^{\omega-1}(x_1)E(x_2) + E^\omega(x_2) + 1](1+E(x_1)+E(x_2))|\xi_s^1 - \xi_s^2\|(dx_1)\mu_s^1(dx_2)$$
$$\leq \kappa(C, \omega, e_0, e_1) \int_0^t ds \|\xi_s^1 - \xi_s^2\|_{1+E^\omega} \|\mu_s^1\|_{1+E^{\omega+1}},$$

and the contribution of the second term by

$$\kappa(C, \omega, e_0, e_1) \int_0^t ds \|\mu_s^1 - \mu_s^2\|_{1+E^{\omega+1}} \|\xi_s^2\|_{1+E^{\omega+1}}$$
$$\leq \kappa(C, \omega, e_0, e_1, (E^{2+\omega}, \mu_0^1 + \mu_0^2))t\|\mu_0^1 - \mu_0^2\|_{1+E^{\omega+1}} \|\xi_0^2\|_{1+E^{\omega+1}}.$$

It remains to apply Gronwall's lemma to complete the proof of statement (iii).

(iv) General results on the derivatives of the evolution systems with respect to the initial data seem not to be applied directly for (1.5). But they can be applied to the cut-off equations

[1] Note that $1+E(x)$ should be of the order $o(1)_{x\to\infty}(\psi_2/\psi_1)(x)$ in order to fulfill the condition on the intensity a_t from Theorem C.1. Hence the necessity of the condition $\omega < \beta - 2$.

(and this is the only reason for introducing these cut-offs in our exposition). Namely, as can be easily seen (this is a simplified "bounded coefficients" version of Proposition 2.2(ii)), the solution μ_t^n to the cut-off version of the kinetic equations (1.5) satisfies this equation strongly in the norm topology of $\mathcal{M}_{1+E^{\beta-\epsilon}}$ for any $\epsilon > 0$. Moreover, μ_n^t depends Lipshtiz continuously on μ_0 in the same topology, the r.h.s. of the cut-off version of (1.5) is differentiable with respect to μ_t in the same topology and ξ_t^n satisfies the equation in variation (4.2) in the same topology. Hence it follows from Proposition 6.5.3 of [25] that

$$\xi_t^n = \xi_t^n(\mu_0; x; dz) = \lim_{s \to 0_+} \frac{1}{s}(\mu_t^n(\mu_0, +s\delta_x) - \mu_t(\mu_0))$$

in the norm topology of $\mathcal{M}_{1+E^{\beta-\epsilon}}$ with $\epsilon > 0$. Consequently

$$(g, \mu_t^n(\mu_0 + h\delta_x)) - (g, \mu_t^n(\mu_0)) = \int_0^h (g, \xi_t^n(\mu_0 + s\delta_x; x; \cdot))\, ds$$

for all $g \in C_{1+E^{\beta-\epsilon},\infty}(X)$ and $\epsilon > 0$. Using statement (ii) and the dominated convergence theorem one deduces that

$$(g, \mu_t(\mu_0 + h\delta_x)) - (g, \mu_t(\mu_0)) = \int_0^h (g, \xi_t(\mu_0 + s\delta_x; x; \cdot))\, ds \tag{4.11}$$

for all $g \in C_{1+E^\gamma}(X)$ with $\gamma < \beta - 2$. Again using the dominated convergence and the fact that ξ_t are bounded in $\mathcal{M}_{1+E^{\beta-1}}$ (as they are $\star$-weak continuous there) one deduces that (4.11) holds for $g \in C_{1+E^{\beta-1},\infty}(X)$. Next, for these g the expression under the integral in the r.h.s. of (4.11) depends continuously on s due to Theorem C.2 (iv), which justifies the weak form of the limit (4.1) (in the $\star$-weak topology of $\mathcal{M}_{1+E^{\beta-1}}$). At last, by statement (iii) ξ_t depends Lipshitz continuously on s in the r.h.s. of (4.11) in the norm topology of $\mathcal{M}_{1+E^\gamma}$ with $\gamma < \beta - 2$. As ξ_t are bounded in $\mathcal{M}_{1+E^{\beta-2}}$ it implies that ξ_t depends continuously on s in the r.h.s. of (4.11) in the norm topology of $\mathcal{M}_{1+E^\gamma}$ with $\gamma < \beta - 1$. Hence (4.11) implies (4.1) in the norm topology of $\mathcal{M}_{1+E^\gamma}(X)$, $\gamma < \beta - 1$, completing the proof of Proposition 4.2.

Proposition 4.3 *(i) Under the assumptions of Proposition 2.1 there exists a unique solution η_t to* (4.4) *in the sense that $\eta_0 = 0$, η_t is a $*$-weakly continuous function $t \mapsto \mathcal{M}_{1+E^{\beta-2}}$ and* (4.4) *holds for $g \in C_{1+E}(X)$. Moreover*

$$\begin{aligned}\|\eta_t(x, w; \cdot)\|_{1+E^\omega} \le\ & \kappa(C, t, \|\mu_0\|_{1+E^\beta}) \\ & \times \sup_{s \in [0,t]} (\|\xi_s(x; \cdot)\|_{1+E^{\omega+\alpha}} \|\xi_s(w; \cdot)\|_{1+E} + \|\xi_s(w; \cdot)\|_{1+E^{\omega+\alpha}} \|\xi_s(x; \cdot)\|_{1+E})\end{aligned} \tag{4.12}$$

for $1 \le \omega \le \beta - 2$ and some κ.

(ii) If η_t^n are defined analogously to η_t but from the cut-off kernels K_n, then $\eta_t^n \to \eta_t$ in the norm topology of $\mathcal{M}_{1+E^\gamma}$ with $\gamma < \beta - 3$ and in the $*$-weak topology of $\mathcal{M}_{1+E^{\beta-2}}$.

(iii) η_t can be defined by the r.h.s. of (4.3) *in the norm topology of $\mathcal{M}_{1+E^\gamma}$ with $\gamma < \beta - 2$ and in the $*$-weak topology of $\mathcal{M}_{1+E^{\beta-2}}$.*

Proof. (i) Linear equation (4.4) differs from equation (4.2) by an additional non homogeneous term. Hence one deduces from Proposition 2.1 (i) the well posedness of this equation and the explicit formula

$$\eta_t(x, w) = \int_0^t V^{t,s} \Omega_s(x, w) ds, \tag{4.13}$$

where $V^{t,s}$ is a resolving operator to the Cauchy problem of equation (4.2) given by Proposition 4.2(i) (or directly form Theorem C.2) and $\Omega_s(x,w)$ is the measure defined weakly as

$$(g, \Omega_s(x,w)) = \int_{X\times X}\int_X (g(y) - g(x_1) - g(x_2))K(x_1, x_2; dy)\xi_t(x; dx_1)\xi_t(w; dx_2). \tag{4.14}$$

From this formula and the properties of ξ_t obtained above statement (i) follows.

(ii) This follows from (4.13) and Proposition 2.2(ii).

(iii) As in the proof of Proposition 2.2(iv), we first prove the formula

$$(g, \xi_t(\mu_0 + h\delta_w; x, \cdot)) - (g, \xi_t(\mu_0; x, \cdot)) = \int_0^h (g, \eta_t(\mu_0 + s\delta_w; x, w; \cdot))ds \tag{4.15}$$

for $g \in C_\infty(X)$ by using the approximation η_t^n, and the dominated convergence. Then the validity of (4.15) is extended to all $g \in C_{1+E^{\beta-2},\infty}$ using the dominated convergence and the above obtained bounds for η_t and ξ_t. By continuity of the expression under the integral in the r.h.s. of (4.14) we justify the limit (4.3) in the $*$-weak topology of $\mathcal{M}_{1+E^{\beta-2}}(X)$ completing the proof of Proposition 4.3.

5 Derivatives with respect to initial data: estimates

Straightforward application of Theorem C.2 of the Appendix would give exponential dependence on (E^β, μ_0) of the constant κ in (4.8). And this is not sufficient for our purposes. The aim of this Section is to obtain more precise estimates for ξ_t. Unlike the rough results of the previous section that can be more or less straightforwardly extended to very general models with fragmentation, collision breakage and their non-binary versions (analyzed in [3], [17], [18]), the arguments of this section use more specific properties of the model under consideration.

We shall use the notations of the previous section, assuming in particular that A_t and B_t are given by (4.9), (4.10) respectively. Due to the results of the previous section we are able to assume that all the Cauchy problems we are dealing with are well-posed. Recall that we denote by $U^{t,r}$ the backward propagator of the equation (2.10).

Let us start with an estimate of the backward propagator $U_A^{t,r}$ of the equation $\dot{g} = -A_t g$ that holds without additional assumptions (C1)-(C3).

Proposition 5.1 *For all $k \geq 0$, $U_A^{t,r}$ is a contraction in $C_{(1+E^k)^{-1}}$ and*

$$|U_A^{t,r} g(x)| \leq \kappa(C, k, r, e_0, e_1)\|g\|_{1+E^k}[(1 + E^k)(x) + (E^{k+1}, \mu_0)]. \tag{5.1}$$

Proof. $U_A^{t,r}$ is a contraction in $C_{(1+E^k)^{-1}}$ by Proposition C.1, because $A_t((1 + E^k)^{-1}) \leq 0$ (and this holds, because $E^k(y) \geq E^k(x)$ in the support of the measure $K(z, x; dy)$). Next

$$A_t(1 + E^k)(x) \leq C\int[(E(x) + E(z))^k - E^k(x)](1 + E(x) + E(z))\mu_t(dz).$$

Using the elementary inequality

$$((a + b)^k - a^k)(1 + a + b) \leq c(k)(a^k(1 + b) + b^{k+1} + 1)$$

that is valid for all positive a, b, k with some constants $c(k)$ yields

$$A_t(1+E^k)(x) \le Cc(k)[E^k(x)(e_0+e_1)+e_0+(E^{k+1},\mu_t)].$$

Then by (2.2)

$$A_t(1+E^k)(x) \le \kappa(C,k,t,e_0,e_1)[E^k(x)+1+(E^{k+1},\mu_0)].$$

Hence (5.1) follows by Lemma C.2 and the fact that $U_A^{t,r}$ is a contraction.

To simplify formulas we shall often use the following elementary inequalities :

$$\begin{aligned} &(a) \qquad (E^l,\nu)(E^k,\nu) \le 2(E^{k+l-1},\nu)(E,\nu), \\ &(b) \qquad (E^k,\nu)E(x) \le (E^{k+1},\nu)+(E,\nu)E^k(x). \end{aligned} \tag{5.2}$$

valid for arbitrary positive ν and $k, l \ge 1$. [2]

Proposition 5.2 *Under condition (C1) suppose $k \ge 1$. Then*

$$|U^{t,r}g(x)| \le \kappa(C,k,r,e_0,e_1)\|g\|_{1+E^k}[1+E^k(x)+(E^{k+1},\mu_0)(1+E(x))], \tag{5.3}$$

$$\sup_{s\le t}\|\xi_s(\mu_0;x,\cdot)\|_{1+E^k} \le \kappa(C,t,e_0,e_1)[1+E^k(x)+(1+E(x))(E^{k+1},\mu_0)], \tag{5.4}$$

and

$$\begin{aligned} \sup_{s\le t}&\|\eta_s(\mu_0;x,w;\cdot)\|_{1+E^k} \le \kappa(C,k,t,e_0,e_1) \\ &\times[(1+E^{k+1}(x)+(E^{k+1},\mu_0)(1+E^2(x))+(E^{k+3},\mu_0)(1+E(x)))(1+E(w)) \\ &+(1+E^{k+1}(w)+(E^{k+1},\mu_0)(1+E^2(w))+(E^{k+3},\mu_0)(1+E(w)))(1+E(x))]. \end{aligned} \tag{5.5}$$

Proof. The simplicity of condition (C1) stems from the observation that the two dimensional functional space generated by the function E and constants is invariant under both A_t and B_t, and also the full image of B_t belongs to this space. Hence representing the solution to $\dot g = -(A_t - B_t)g$ as

$$g = U_A^{t,r}g_r + \tilde g \tag{5.6}$$

one finds that $\tilde g$ belongs to the above mentioned two dimensional space and satisfies the equation

$$\dot{\tilde g} = -(A_t-B_t)\tilde g + B_tU_A^{t,r}g_r, \qquad \tilde g\mid_{t=r}=0, \qquad t\le r. \tag{5.7}$$

The corresponding homogeneous Cauchy problem

$$\dot\phi = -(A_t-B_t)\phi, \qquad \phi_r = \alpha+\beta E,$$

can be written as

$$\dot\alpha_t + \dot\beta_t E(x) = C\alpha_t(e_1+(1,\mu_t)E(x)), \quad \alpha_r=\alpha, \beta_r=\beta$$

[2]Say, to get (a), one writes $(E^l,\nu)(E^k,\nu) = \int\int E^l(x)E^k(y)\nu(dx)\nu(dy)$ and decomposes this integral into the sum or two integrals over the domains $\{E(x)\ge E(y)\}$ and otherwise; then, say, the first integral is estimated by $\int\int E^{l+k-1}(x)E(y)\nu(dx)\nu(dy)$

in terms of $\phi = \alpha_t + \beta_t E(x)$ and clearly solves explicitly as

$$\phi_t = \alpha e^{-e_1(r-t)} + \left[\beta + \alpha \int_t^r (1, \mu_s) e^{-e_1(r-s)} ds\right] E(x),$$

which implies that

$$\|\phi_t\|_{1+E} \leq \kappa(r, e_0) \|\phi_r\|_{1+E}.$$

It follows from (5.1) that

$$\begin{aligned} |B_t U_A^{t,r} g_r(x)| &\leq \kappa(C, r, e_0, e_1) \|g_r\|_{1+E^k} [B_t(1 + E^k) + (E^{k+1}, \mu_0) B_t 1](x) \\ &\leq \kappa(C, r, e_0, e_1) \|g_r\|_{1+E^k} (1 + (E^{k+1}, \mu_0))(1 + E(x)). \end{aligned} \tag{5.8}$$

Solving the non-homogeneous equation (5.7) by the Du Hamel principle and using the representation (5.6) yields (5.3). But by duality one gets

$$\|\xi_s(\mu_0; x, \cdot)\|_{1+E^k} = \sup\{(g, \xi_s(\mu_0; x, \cdot)) : \|g\|_{1+E^k} \leq 1\}$$
$$= \sup\{(U^{0,s} g, \delta_x) : \|g\|_{1+E^k} \leq 1\} = \sup\{U^{0,s} g(x) : \|g\|_{1+E^k} \leq 1\},$$

which implies (5.4).

Now from (4.13)

$$\begin{aligned} \sup_{s \leq t} \|\eta(\mu_0; x, w; \cdot)\|_{1+E^k} &\leq t \sup_{s \leq t} \|V^{t,s} \Omega_s(x, w)\|_{1+E^k} = t \sup_{s \leq t} \sup_{|g| \leq 1+E^k} (U^{s,t} g, \Omega_s(x, w)) \\ &\leq \kappa(C, t, e_0, e_1) \sup_{s \leq t} \sup\{(g, \Omega_s(x, w)) : |g(y)| \leq 1 + E^k(y) + (1 + E(y))(E^{k+1}, \mu_0)\} \\ &\leq \kappa(C, k, t, e_0, e_1) \sup_{s \leq t} \int\int [1 + E^k(x_1) + E^k(x_2) + (E^{k+1}, \mu_0)(1 + E(x_1) + E(x_2))] \\ &\quad (1 + E(x_1) + E(x_2)) \xi_s(x; dx_1) \xi_s(w; dx_2). \end{aligned}$$

Dividing this integral into two parts with $E(x_1) \geq E(x_2)$ and $E(x_1) \leq E(x_2)$ one can estimate the first part as

$$\begin{aligned} &\kappa \sup_{s \leq t} \int\int [1 + E^k(x_1) + (E^{k+1}, \mu_0)(1 + E(x_1))](1 + E(x_1)) \xi_s(x; dx_1) \xi_s(w; dx_2) \\ &\leq \kappa \sup_{s \leq t} \|\xi_s(w; \cdot)\| \left(\|\xi_s(x; \cdot)\|_{1+E^{k+1}} + (E^{k+1}, \mu_0) \|\xi_s(x; \cdot)\|_{1+E^2}\right) \\ &\leq \kappa(1 + E(w))[1 + E^{k+1}(x) + (1 + E(x))(E^{k+2}, \mu_0) \\ &\quad + (E^{k+1}, \mu_0)(1 + E^2(x) + (1 + E(x))(E^3, \mu_0)] \\ &\leq \kappa(1 + E(w))[1 + E^{k+1}(x) + (E^{k+1}, \mu_0)(1 + E^2(x)) + (E^{k+3}, \mu_0)(1 + E(x))], \end{aligned}$$

where we used both (5.2)(a) and (5.2)(b). As the integral over the second part is estimated similarly one arrives at (5.5).

Proposition 5.3 *Under condition (C2)*

$$\|U^{t,r}\|_{C_{1+\sqrt{E}}} \leq \exp\{4C(t - r)(e_0 + e_1)\} \tag{5.9}$$

and the estimates (5.3)-(5.5) *hold for all* $k \geq 1$.

Proof. Since

$$A_t(1+\sqrt{E})(z) = \int\int (\sqrt{E(z)+E(x)} - \sqrt{E(z)})K(z,x;dy)\mu_t(dx)$$
$$\le C\int_X \sqrt{E(x)}(1+\sqrt{E(z)})(1+\sqrt{E(x)})\mu_t(dx)$$
$$\le C(1+\sqrt{E(z)})(\sqrt{E}+E,\mu_t) \le C(e_0+2e_1)(1+\sqrt{E(z)}),$$

according to Proposition C.1 the positivity preserving backward propagator $U_A^{r,t}$ of the equation $\dot{g} = -A_t g$ is bounded in $C_{1+\sqrt{E}}(X)$ with the norm not exceeding $\exp\{C(t-r)(e_0+2e_1)\}$. On the other hand

$$B_t(1+\sqrt{E})(z) \le C\int (1+\sqrt{E(x)})^2(1+\sqrt{E(z)})\mu_t(dx)$$
$$\le 2C(e_0+e_1)(1+\sqrt{E(z)}).$$

Hence B_t are uniformly bonded in $C_{1+\sqrt{E}}(X)$ with the norm not exceeding $2C(e_0+e_1)$. Hence (5.9) follows from the series representation (C.16) for the backward propagator $U^{r,t}$ of the equation $\dot{g} = -(A_t - B_t)g$.

Now we use the same arguments as in the proof of Proposition 5.2 with

$$|B_t U_A^{t,r} g_r(x)| \le \kappa(C, t-r, e_0, e_1)\|g_r\|_{1+E^k}(1+\sqrt{E(x)})(1+(E^{k+1},\mu_0))$$

instead of (5.8). Namely, in the representation of the solutions to $\dot{g} = -(A_t - B_t)g$ by the series (C.16) the first term is independent of B_t and all other terms belong to $C_{1+\sqrt{E}}(X)$ and applying the above estimates for $U_A^{t,r}$ and B_t in this space one deduces (5.3). Other estimate follows now straightforwardly as in the previous Proposition (even with some improvements that we do not take into account).

Proposition 5.4 *Under condition (C3) for any $k \ge 0$ the spaces $C^{1,0}_{1+E^k}$ and $C^{2,0}_{1+E^k}$ (see Appendix A for these notations) are invariant under $U^{t,r}$ and*

$$
\begin{aligned}
&(a) \qquad |(U^{t,r}g)'(x)| \le \kappa(C,r,k,e_0,e_1)\|g\|_{C^{1,0}_{1+E^k}}(1+E^k(x)+(E^{k+1},\mu_0)),\\
&(b) \qquad |(U^{t,r}g)''(x)| \le \kappa(C,r,k,e_0,e_1)\|g\|_{C^{2,0}_{1+E^k}}(1+E^k(x)+(E^{k+1},\mu_0)),
\end{aligned}
\tag{5.10}
$$

$$\sup_{s\le t}\|\xi_s(\mu_0;x;\cdot)\|_{\mathcal{M}^1_{1+E^k}} \le \kappa(C,r,k,e_0,e_1)[E(x)(1+(E^{k+1},\mu_0))+E^{k+1}(x)], \tag{5.11}$$

and

$$
\begin{aligned}
\sup_{s\le t}&\|\eta_s(\mu_0;x,w;,\cdot)\|_{\mathcal{M}^2_{1+E^k}} \le \kappa(C,t,k,e_0,e_1)(1+(E^{k+1},\mu_0))\\
&\times[(E(x)(1+E^{k+2},\mu_0)+E^{k+2}(x))E(w)+(E(w)(1+E^{k+2},\mu_0)+E^{k+2}(w))E(x)].
\end{aligned}
\tag{5.12}
$$

Proof. Notice first that if $g_r(0) = 0$, then $g_t = 0$ for all t according to the evolution described by the equation $\dot{g} = -(A_t - B_t)g$. Hence the space of functions vanishing at the origin is invariant under this evolution.

Recall that $E(x) = x$ in case (C3). Differentiating the equation $\dot{g} = -(A_t - B_t)g$ with respect to the space variable x leads to the equation

$$\dot{g}'(x) = -A_t(g')(x) - \int (g(x+z) - g(x) - g(z)) \frac{\partial K}{\partial x}(x,z)\mu_t(dz). \tag{5.13}$$

For functions g vanishing at the origin this can be rewritten as

$$\dot{g}'(x) = -A_t g' - D_t g'$$

with

$$D_t\phi(x) = \int \left(\int_x^{x+z} \phi(y)dy - \int_0^z \phi(y)dy \right) \frac{\partial K}{\partial x}(x,z)\mu_t(dz).$$

Since

$$\|D_t\phi\| \leq 2C\|\phi\|(E,\mu_t) = 2Ce_1\|\phi\|,$$

and $U_A^{t,r}$ is a contraction, it follows from representation (C.16) with D_t instead of B_t that

$$\|U^{t,r}\|_{C_1^{1,0}(X)} \leq \kappa(C, r-t, e_0, e_1),$$

proving (5.10)(a) for $k = 0$. Next, for $k > 0$

$$\begin{aligned}|D_t\phi(x)| &\leq C\|\phi\|_{1+E^k} \int ((x+z)^{k+1} - x^{k+1} + z^{k+1} + z)\mu_t(dz) \\ &\leq Cc(k)\|\phi\|_{1+E^k} \int (x^k z + z^{k+1} + z)\mu_t(dz),\end{aligned}$$

which by (2.2) does not exceed

$$c(C,k,e_1)\|\phi\|_{1+E^k}[(1+x^k) + (E^{k+1}, \mu_0)].$$

Hence by Proposition 5.1

$$\int_t^r |U_A^{t,s} D_s U_A^{s,r} g(x)| ds \leq (r-t)\kappa(C,r,k,e_0,e_1)\|g\|_{1+E^k}[1 + E^k(x) + (E^{k+1},\mu_0)],$$

which by induction implies

$$\begin{aligned}&\int_{t \leq s_1 \leq \cdots \leq s_n \leq r} |U_A^{t,s_1} D_{s_1} \cdots D_{s_n} U_A^{s_n,r} g(x)| ds_1 \cdots ds_n \\ &\qquad \leq \frac{(r-t)^n}{n!} \kappa^n(C,r,k,e_0,e_1)\|g\|_{1+E^k}[1 + E^k(x) + (E^{k+1},\mu_0)].\end{aligned}$$

Hence (5.10)(a) follows from the representation (C.16) to the solution of (5.13).

Differentiating (5.13) leads to the equation

$$\dot{g}''(x) = -A_t(g'')(x) - \psi_t, \tag{5.14}$$

where

$$\psi_t = 2\int (g'(x+z)-g'(x))\frac{\partial K}{\partial x}(x,z)\mu_t(dz) + \int \left(\int_x^{x+z} g'(y)dy - \int_0^z g'(y)dy\right)\frac{\partial^2 K}{\partial x^2}(x,z)\mu_t(dz).$$

We know already that for $g_r \in C^2_{1+E^k}$ the function g' belongs to $1+E^k$ with the bound given by (5.10)(a). Hence by the Du Hamel principle the solution to (5.14) can be represented as

$$g_t'' = U_A^{t,r} g_r'' + \int_t^r U_A^{t,s}\psi_s ds.$$

As

$$|\psi_t(x)| \le \kappa(C, r-t, e_0, e_1)(1+E^k(x)+(E^{k+1},\mu_0)),$$

(5.10)(b) follows, completing the proof of (5.10), which by duality implies (5.11).

Next, arguing as in the proof of Proposition 5.2 one gets

$$\sup_{s\le t}\|\eta_s(\mu_0;x,w;.)\|_{\mathcal{M}^2_{1+E^k}} \le t\sup_{s\le t}\sup\{|(U^{s,t}g,\Omega_s(x,w))| : \|g\|_{C^{2,0}_{1+E^k}} \le 1\}$$
$$\le \kappa(C,t,e_0,e_1)\sup_{s\le t}\sup_{g\in\Pi_k}(g,\Omega_s(x,w)),$$

where

$$\Pi_k = \{g : g(0)=0, \max(|g'(y)|,|g''(y)|) \le 1+E^k(y)+(E^{k+1},\mu_0)\}.$$

It is convenient to introduce a two times continuously differentiable function χ on $\mathbf{R}$ such that $\chi(x)\in[0,1]$ for all x, and $\chi(x)$ equals one or zero respectively for $x\ge 1$ and $x\le -1$. Then write $\Omega_s = \Omega_s^1+\Omega_s^2$ with Ω^1 (respectively Ω^2) being obtained by (4.14) with $\chi(x_1-x_2)K(x_1,x_2)$ (respectively $(1-\chi(x_1-x_2))K(x_1,x_2)$) instead of $K(x_1,x_2)$. If $g\in\Pi_k$, one has

$$(g,\Omega_s^1(x,w)) = \int\int (g(x_1+x_2)-g(x_1)-g(x_2))\chi(x_1-x_2)K(x_1,x_2)\xi_s(x;dx_1)\xi_s(w;dx_2),$$

which is bounded in magnitude by

$$\|\xi_s(w,\cdot)\|_{\mathcal{M}^1_1(X)}\sup_{x_2}\left|\frac{\partial}{\partial x_2}\int[(g(x_1+x_2)-g(x_1)-g(x_2))\chi(x_1-x_2)K(x_1,x_2)]\xi_s(x;dx_1)\right|$$
$$\le \|\xi_s(w,\cdot)\|_{\mathcal{M}^1_1(X)}\|\xi_s(x,\cdot)\|_{\mathcal{M}^1_{1+E^{k+1}}(X)}$$
$$\times\sup_{x_1,x_2}\left|(1+E^{k+1}(x_1))^{-1}\frac{\partial^2}{\partial x_2\partial x_1}[(g(x_1+x_2)-g(x_1)-g(x_2))\chi(x_1-x_2)K(x_1,x_2)]\right|.$$

Since

$$\frac{\partial^2}{\partial x_2\partial x_1}[(g(x_1+x_2)-g(x_1)-g(x_2))\chi(x_1-x_2)K(x_1,x_2)]$$
$$= g''(x_1+x_2)(\chi K)(x_1,x_2) + (g'(x_1+x_2)-g'(x_2))\frac{\partial(\chi K)(x_1,x_2)}{\partial x_1}$$
$$+ (g'(x_1+x_2)-g'(x_1))\frac{\partial(\chi K)(x_1,x_2)}{\partial x_2} + (g(x_1+x_2)-g(x_1)-g(x_2))\frac{\partial^2(\chi K)(x_1,x_2)}{\partial x_1\partial x_2},$$

this expression does not exceed in magnitude $C(1+E^{k+1}(x_1)+(E^{k+1},\mu_0)(1+E(x_1))$ (up to a constant multiplier). Consequently

$$|(g,\Omega_s^1(x,w)|\le\kappa(C)\|\xi_t(w,\cdot)\|_{\mathcal{M}_1^1(X)}\|\xi_t(x,\cdot)\|_{\mathcal{M}^1_{1+E^{k+1}}(X)}(1+(E^{k+1},\mu_0)).$$

Of course, the norm of Ω_s^2 is estimated in the same way. Consequently (5.11) leads to (5.12) and completes the proof of Proposition 5.4.

We shall prove now the Lipschitz continuity of the solutions to our kinetic equation with respect to initial data in the norm-topology of the space $\mathcal{M}^1_{1+E^k}$.

Proposition 5.5 *Under the condition (C3) for $k\ge 0$ and $m=1,2$*

$$\sup_{s\le t}\|\mu_s(\mu_0^1)-\mu_s(\mu_0^2)\|_{\mathcal{M}^m_{1+E^k}}\le\kappa(C,t,k,e_0,e_1)(1+E^{1+k},\mu_0^1+\mu_0^2)\|\mu_0^1-\mu_0^2\|_{\mathcal{M}^m_{1+E^k}}\tag{5.15}$$

Proof. By (4.1) and (B.1)

$$(g,\mu_t(\mu_0^1)-\mu_t(\mu_0^2))=\int_0^t ds\int\int g(y)\xi_t(\mu_0^2+s(\mu_0^1-\mu_0^2);x;dy)(\mu_0^1-\mu_0^2)(dx).\tag{5.16}$$

Since

$$(g,\xi_t(Y;x;.))=(U^{0,t}g,\xi_0(Y,x;.))=(U^{0,t}g)(x),$$

it follows from Proposition 5.4 that $(g,\xi_t(Y;x;.))$ belongs to $C^{m,0}_{1+E^k}$ as a function of x whenever g belongs to this space and that

$$\|(g,\xi_t(Y;x;.))\|_{C^{m,0}_{1+E^k}(X)}\le\kappa(C,t,k,e_0,e_1)\|g\|_{C^{m,0}_{1+E^k}(X)}(1+(E^{k+1},Y)).$$

Consequently (5.15) follows from (5.16).

We shall discuss now the L^2-version of our estimates.

Proposition 5.6 *Under condition (C3) assume f is a positive either non-decreasing or bounded function. Then $U_A^{t,r}$ are contractions in $L_{2,1/f}$. (Thus $U_A^{t,r}$ yield natural examples of sub-Markovian propagators with growing coefficients.)*

Proof. First observe that

$$\int_0^\infty(u(x+y)-u(x))g^2(x)u(x)dx\le 0\tag{5.17}$$

for any $y\ge 0$ and a non-decreasing non-negative g (and any u, if only the integral is well defined). In fact, it is equivalent to

$$(T_yu,u)_{L_{2,1/g}}\le(u,u)_{L_{2,1/g}},$$

where $T_yu(x)=u(x+y)$, which in turn follows (by Cauchy inequality) from $(T_yu,T_yu)_{L_{2,1/g}}\le(u,u)_{L_{2,1/g}}$. And the latter holds, because

$$(T_yu,T_yu)_{L_{2,1/g}}=\int_0^\infty u^2(x+y)g^2(x)\,dx$$

$$= \int_y^\infty u^2(z)g^2(z-y)\,dz \le \int_y^\infty u^2(x)g^2(x)\,dx.$$

Assume now that f is non-decreasing (and positive). From (5.17) it follows that for arbitrary $y > 0$

$$\int_0^\infty (u(x+y) - u(x))K(x,y)f^2(x)u(x)\,dx \le 0 \tag{5.18}$$

(we used here the assumed monotonicity of the kernel K), and hence $(A_t u, u)_{L_{2,1/f}} \le 0$. Hence $U_A^{t,r}$ can not increase the norm of $L_{2,1/f}$. To conclude that it is actually a semigroup of contractions it remains to observe that due to Proposition 5.1 there exists a dense subspace in $L_{2,1/f}$ that is invariant under $U_A^{t,r}$. Assume now that f is bounded. We again have to show the validity of (5.18) for a dense invariant subspace of functions u. First note that as the evolution $U_A^{t,r}$ is well defined on continuous functions and preserves positivity and differentiability (by Propositions 5.1 and 5.4) it is suffice to show (5.18) for positive functions u with bounded variation. Assuming that this is the case one can represent positive u as the difference $u = u^+ - u^-$ of two positive non-decreasing functions (by decomposing its derivative in its positive and negative parts). As $-(u^-(x+y) - u^-(x)) \le 0$, to show (5.18) one needs to show that

$$\int_0^\infty (u^+(x+y) - u^+(x))K(x,y)f^2(x)u(x)\,dx \le 0,$$

and as $-u^-$ is negative this in turns follows from

$$I_y = \int_0^\infty (u^+(x+y) - u^+(x))K(x,y)f^2(x)u^+(x)\,dx \le 0.$$

Denoting by M an upper bound for f^2 we can write

$$I_y = \int_0^\infty (u^+(x+y) - u^+(x))K(x,y)Mu(x)\,dx - \int_0^\infty (u^+(x+y) - u^+(x))K(x,y)(M - f^2(x))u(x)\,dx,$$

which is negative, because the integrand in the second term is positive and the first term is negative by (5.17).

We shall consider now equation (5.13) that can be written in the form

$$\dot{g}'(x) = -A_t(g')(x) - D_t^1 g' - D_t^2 g', \tag{5.19}$$

where

$$(D_t^1\phi)(x) = \int_0^x \left(\int_x^{x+z} \phi(y)\,dy \frac{\partial K}{\partial x}(x,z) \right) \mu_t(dz),$$

$$(D_t^2\phi)(x) = \int_0^\infty \left(\mathbf{1}_{x<z} \int_x^{x+z} \phi(y)\,dy - \int_0^z \phi(y)\,dy \right) \frac{\partial K}{\partial x}(x,z)\mu_t(dz).$$

Proposition 5.7 *Under condition (C3) for any $f_k(x) = 1 + x^k$, $k > 1/2$, the spaces $L_{2,f_k}^{m,0}$, $m = 1,2$, are invariant under $U^{t,r}$ and*

$$\|U^{t,r}\|_{L_{2,f_k}^{m,0}} \le \kappa(C,r,k,e_0,e_1)(1 + (E^{k+1/2},\mu_0)), \quad m = 1,2.$$

Moreover, for $g \in L^{m,0}_{2,f_k}$ one can represent $(U^{t,r}g)'$ as the sum of a function from $L^{m,0}_{2,f_k}$ with the norm not exceeding $\kappa(C,r,k,e_0,e_1)\|g\|_{L^{m,0}_{2,f_k}}$ and a uniformly bounded function with the sup-norm not exceeding $\kappa(C,r,k,e_0,e_1)\|g\|_{L^{m,0}_{2,f_k}}(1+(E^{k+1/2},\mu_0))$. Consequently

$$\sup_{s\le t}\|\xi_s(\mu_0;x;\cdot)\|_{(L^{1,0}_{2,f_k})'} \le \kappa(C,r,k,e_0,e_1)[E(x)(E^{k+1},\mu_0)+(1+E^{k+1/2}(x)]. \tag{5.20}$$

Proof. Let us show first that

$$\|D^1_t\|_{L_{2,f_k}} \le e_1 2^k C. \tag{5.21}$$

In fact, for a continuous positive ϕ and an arbitrary $z>0$

$$\|\mathbf{1}_{z\le x}\int_x^{x+z}\phi(y)\,dy\|^2_{L_{2,f_k}} = \lim_{n\to\infty}\int \mathbf{1}_{z\le x}\sum_{i,j=1}^n \phi(x+\frac{jz}{n})\phi(x+\frac{iz}{n})\frac{z^2}{n^2}f_k^{-2}(x)\,dx$$

$$\le z^2\lim_{n\to\infty}\int \mathbf{1}_{z\le x}\sum_{i,j=1}^n(\phi/f_k)(x+\frac{jz}{n})(\phi/f_k)(x+\frac{iz}{n})\frac{1}{n^2}2^{2k}\,dx,$$

because

$$\frac{1}{f_k(x)} \le \frac{2^k}{f_k(2x)} \le \frac{2^k}{f_k(x+jz/n)}$$

for all $j\le n, z\le x$. Taking now into account that

$$\int \mathbf{1}_{z\le x}\sum_{i,j=1}^n(\phi/f_k)(x+\frac{jz}{n})(\phi/f_k)(x+\frac{iz}{n})\,dx \le \|\phi/f_k\|^2_{L_2}$$

one deduces that

$$\|\mathbf{1}_{z\le x}\int_x^{x+z}\phi(y)\,dy\|^2_{L_{2,f_k}} \le z^2 2^{2k}\|\phi\|^2_{L_{2,f_k}}.$$

Consequently

$$\|D^1_t\phi\|_{L_{2,f_k}} \le C\int_0^\infty \|\mathbf{1}_{z\le x}\int_x^{x+z}\phi(y)\,dy\|_{L_{2,f_k}}\mu_t(dz) \le C2^k e_1\|\phi\|_{L_{2,f_k}},$$

which implies (5.21).

As clearly the same bounds hold for $\|D^1_t\|_{C(\mathbf{R}_+)}$, the equation

$$\dot g'(x) = -A_t(g')(x) - D^1_t g'$$

specifies a propagator $\tilde U^{t,r}$, $t\in[0,r]$, of bounded operators in both $C(\mathbf{R}_+)$ and $L_{2,f_k}(\mathbf{R}_+)$ with uniform bounds depending on r,k,e_0,e_1. Next

$$|(D^2_t\phi)(x)| \le 2C\int\int_0^{2z}|\phi(y)|dy\mu_t(dz)$$

for all x, which by Cauchy-Schwartz inequality does not exceed

$$2C\|\phi\|_{L_{2,f_k}}\int\sqrt{\int_0^{2z}f_k^2(y)\,dy}\mu_t(dz) \le Cc(k)\|\phi\|_{L_{2,f_k}}(1+(E^{k+1/2},\mu_0)).$$

Hence writing the solution to the Cauchy problem for equation 5.13 with $\phi_r \in L_{2,f_k}$ as a perturbation series (C.16) with respect to perturbation D_t^2 one represents the solution as the sum $\phi_1^t + \phi_2^t$ with

$$\|\phi_1^t\|_{L_{2,f_k}} \leq c(k, e_0, e_1, r)\|\phi_r\|_{L_{2,f_k}}$$

and

$$\|\phi_2^t\|_{C(\mathbf{R}_+)} \leq (1 + (E^{k+1/2}, \mu_0))\kappa(C, k, e_0, e_1, r)\|\phi_r\|_{L_{2,f_k}}$$

so that

$$\|\phi_2^t\|_{L_{2,f_k}} \leq (1 + (E^{k+1/2}, \mu_0))\kappa(C, k, e_0, e_1, r)\|\phi_r\|_{L_{2,f_k}}$$

whenever $k > 1/2$. In particular for these k

$$\|U^{t,r}\|_{L^{1,0}_{2,f_k}} \leq (1 + (E^{k+1/2}, \mu_0))\kappa(C, k, e_0, e_1, r).$$

As $(\xi_s(\mu_0; x; .), g) = (\delta_x, U^{0,s}g)$, this implies (5.20) by (A.1). The evolution $U^{t,r}$ in the space $L^{2,0}_{2,f_k}$ is analyzed quite similarly.

We conclude with the following analog of Proposition 5.5, whose proof follows from Proposition 5.7 by the same argument as Proposition 5.5 follows from Proposition 5.4.

Proposition 5.8 *Under the condition (C3) for $k > 1/2$ and $m = 1, 2$*

$$\sup_{s \leq t} \|\mu_s(\mu_0^1) - \mu_s(\mu_0^2)\|_{(L^{m,0}_{2,f_k})'} \leq \kappa(C, t, k, e_0, e_1)(1 + E^{1+k/2}, \mu_0^1 + \mu_0^2)\|\mu_0^1 - \mu_0^2\|_{(L^{m,0}_{2,f_k})'} \qquad (5.22)$$

6 The rate of convergence in the LLN

Proof of Theorem 2.1. Recall that $\mu_t(Y)$ means the solution to equation (1.5) with initial data $\mu_0 = Y$ given by Proposition 2.1 with a $\beta \geq 2$. We shall write shortly $Y_t = \mu_t(Y)$ so that $T_t F(Y) = F(\mu_t(Y)) = F(Y_t)$. For a function $F(Y) = (g, Y^{\otimes m})$ with $g \in C^{sym}_{(1+E)^{\otimes m},\infty}(X^m)$, $m \geq 1$, and $Y = h\delta_{\mathbf{x}}$ one has

$$T_t F(Y) - T_t^h F(Y) = \int_0^t T_{t-s}^h (L_h - \mathcal{L}) T_s F(Y) \, ds. \qquad (6.1)$$

As $T_t F(Y) = (g, Y_t^{\otimes m})$, Propositions 4.2 and 4.3 yield

$$\delta T_t F(Y; x) = m \int_{X^m} g(y_1, y_2, \ldots, y_m) \xi_t(Y; x; dy_1) Y_t^{\otimes(m-1)}(dy_2 \cdots dy_m),$$

and

$$\delta^2 T_t F(Y; x, w) = m \int_{X^m} g(y_1, y_2, \ldots, y_m) \eta_t(Y; x, w; dy_1) Y_t^{\otimes(m-1)}(dy_2 \cdots dy_m)$$
$$+ m(m-1) \int_{X^m} g(y_1, y_2, \ldots, y_m) \xi_t(Y; x; dy_1) \xi_t(Y; w; dy_2) Y_t^{\otimes(m-2)}(dy_3 \cdots dy_m). \qquad (6.2)$$

Let us estimate the difference $(L_h - \mathcal{L}) T_t F(Y)$ using (3.5) (with $T_t F$ instead of F). Let us analyze only the more weird second term in (3.5), as the first one is analyzed similar, but much simpler. We are going to estimate separately the contribution to the last term of (3.5) of the first and second term in (6.2).

Assume that the condition (C1) or (C2) holds and a $k \geq 1$ is chosen. Note that the norm and the first moment $(E, \cdot)$ of $Y + sh(\delta_y - \delta_{x_i} - \delta_{x_j})$ do not exceed respectively the norm and the first moment of Y. Moreover, for $s \in [0,1]$, $h > 0$ and $x_i, x_j, y \in X$ with $E(y) = E(x_i) + E(x_j)$ one has

$$\begin{aligned}(E^k, Y + sh(\delta_y - \delta_{x_i} - \delta_{x_j})) &= (E^k, Y) + sh(E(x_i) + E(x_j))^k - hE^k(x_i) - hE^k(x_j) \\ &\leq (E^k, Y) + hc(k)(E^{k-1}(x_i)E(x_j) + E(x_i)E^{k-1}(x_j))\end{aligned}$$

with a constant $c(k)$ depending only on k. Consequently by Proposition 5.2

$$\begin{aligned}&\|\eta_t(Y + sh(\delta_y - \delta_{x_i} - \delta_{x_j}); x, w; \cdot)\|_{1+E^k} \leq \kappa(C, k, t, e_0, e_1)(1 + E(w)) \\ &\{1 + E^{k+1}(x) + [(E^{k+1}, Y) + hc(k)(E^k(x_i)E(x_j) + E(x_i)E^k(x_j))](1 + E^2(x)) \\ &+ [(E^{k+3}, Y) + hc(k)(E^{k+2}(x_i)E(x_j) + E(x_i)E^{k+2}(x_j)](1 + E(x))\} + ...,\end{aligned}$$

where by dots is denoted the similar term with x and w interchanging their places. Hence the contribution to the last term of (3.5) of the first term in (6.2) does not exceed

$$\begin{aligned}&\kappa(C, t, k, m, e_0, e_1)\|g\|_{(1+E^k)^{\otimes m}}(1 + E^k, Y)^{m-1}h^3 \sum_{i \neq j}(1 + E(x_i) + E(x_j))^2\{1 + E^{k+1}(x_i) + E^{k+1}(x_j) \\ &+ [(E^{k+1}, Y) + hc(k)(E^k(x_i)E(x_j) + E(x_i)E^k(x_j))](1 + E^2(x_i) + E^2(x_j)) \\ &+ [(E^{k+3}, Y) + hc(k)(E^{k+2}(x_i)E(x_j) + E(x_i)E^{k+2}(x_j)](1 + E(x_i) + E(x_j))\}.\end{aligned}$$

Dividing this sum into two parts, where $E(x_i) \geq E(x_j)$ and respectively vice versa, and noting that by the symmetry it is enough to estimate only the first part, allows to estimate the contribution to the last term of (3.5) of the first term from (6.2) by

$$\begin{aligned}&\kappa(C, t, k, m, e_0, e_1)\|g\|_{(1+E^k)^{\otimes m}}(1 + E^k, Y)^{m-1}h^3 \sum_{i \neq j} \\ &\{1 + E^{k+3}(x_i) + (1 + E^4(x_i))[(E^{k+1}, Y) + hc(k)E^k(x_i)E(x_j)] \\ &+ (1 + E^3(x_i))[(E^{k+3}, Y) + hc(k)E^{k+2}(x_i)E(x_j)]\}.\end{aligned}$$

The main term in this expression (obtained by ignoring the terms with $hc(k)$) is estimated by

$$\kappa\|g\|_{(1+E^k)^{\otimes m}}(1 + E^k, Y)^{m-1}h[(1 + E^{k+3}, Y) + (1 + E^4, Y)(E^{k+1}, Y) + (1 + E^3, Y)(E^{k+3}, Y)],$$

where the first two terms in the square bracket can be estimated by the last one, because

$$(E^4, Y)(E^{k+1}, Y) \leq 2(E^2, Y)(E^{k+3}, Y).$$

It remains to observe that the terms with $hc(k)$ are actually subject to the same bound, as for instance

$$h^4 \sum_{i \neq j} E^k(x_i)E(x_j)(1 + E^4(x_i)) \leq h^2(E^k + E^{k+4}, Y)(E, Y) \leq c(k)h(E^{k+3}, Y)(E, Y)^2.$$

Consequently, the contribution to the last term of (3.5) of the first term in (6.2) does not exceed

$$h\kappa(C,t,k,m,e_0,e_1)\|g\|_{(1+E^k)^{\otimes m}}(1+E^k,Y)^{m-1}(1+E^{k+3},Y)(1+E^3,Y). \tag{6.3}$$

Turning to the contribution of the second term from (6.2) observe that again by Proposition 5.2

$$\|\xi_t(Y+sh(\delta_y-\delta_{x_i}-\delta_{x_j});x;\cdot)\|_{1+E^k} \le \kappa(C,k,t,e_0,e_1)$$
$$\{1+E^k(x)+(1+E(x)[(E^{k+1},Y)+hc(k)(E^k(x_i)E(x_j)+E(x_i)E^k(x_j))]\},$$

so that the contribution of the second term from (6.2) does not exceed

$$\kappa(C,t,k,m,e_0,e_1)\|g\|_{(1+E^k)^{\otimes m}}(1+E^k,Y)^{m-2}h^3\sum_{i\neq j}(1+E(x_i)+E(x_j))\{1+E^k(x_i)+E^k(x_j)$$
$$+(1+E(x_i)+E(x_j))[(E^{k+1},Y)+hc(k)(E^k(x_i)E(x_j)+E(x_i)E^k(x_j))]\}^2,$$

which again by dividing this sum into two parts, where $E(x_i)\ge E(x_j)$ and respectively vice versa, reduces to

$$\kappa(C,t,k,m,e_0,e_1)\|g\|_{(1+E^k)^{\otimes m}}(1+E^k,Y)^{m-2}h^3\sum_{i\neq j}(1+E(x_i))\{1+E^k(x_i)$$
$$+(1+E(x_i)+E(x_j))[(E^{k+1},Y)+hc(k)E^k(x_i)E(x_j)]\}^2.$$

This is again estimated by (6.3). It follows now from (6.1) and Proposition 3.1 that

$$\|T_tF-T_t^hF\|_{C_{(1+E^{k+3},\cdot)(1+E^3,\cdot)(1+E^k,\cdot)^{m-1}}(\mathcal{M}_{h\delta}^{e_0e_1}(X))} \le h\kappa(C,t,k,m,e_0,e_1)\|g\|_{(1+E^k)^{\otimes m}},$$

which is the same as (2.5). The proof of (2.6) is quite the same. It only uses Proposition 5.4 instead of Proposition 5.2.

7 Auxiliary Estimates

The main technical ingredient in the proof of a weak form of CLT (convergence for fixed times, stated in Theorems 2.2- 2.4) is given by the following corollary to Theorem 2.1.

Proposition 7.1 *Under the assumptions of Theorem 2.1 let g_2 be a symmetric continuous function on X^2. Then for any $k\ge 1$*

$$\sup_{s\le t}\left|\mathbf{E}\left(g_2,\left(\frac{Z_s^h(Z_0^h)-\mu_s(\mu_0)}{\sqrt{h}}\right)^{\otimes 2}\right)\right| = \sup_{s\le t}|\mathbf{E}(g_2,(F_s^h(Z_0^h,\mu_0))^{\otimes 2})| = \sup_{s\le t}|(U_{fl}^{h;0,s}(g_2,.))(F_0^h)| \tag{7.1}$$

does not exceed the expression

$$\kappa(C,t,k,e_0,e_1)\|g_2\|_{(1+E^k)^{\otimes 2}(X^2)}(1+(E^{2k+2},Z_0^h+\mu_0))(1+(E^k,Z_0^h+\mu_0))\left(1+\left\|\frac{Z_0^h-\mu_0}{\sqrt{h}}\right\|_{1+E^k}^2\right)$$

for any $k \geq 1$ under the condition (C1) or (C2) and the expression

$$\kappa(C,t,k,e_0,e_1)\|g_2\|_{C^{2,sym}_{(1+E^k)^{\otimes 2}(X^2)}}(1+(E^{k+4},Z_0^h+\mu_0))^3\left(1+\left\|\frac{Z_0^h-\mu_0}{\sqrt{h}}\right\|^2_{\mathcal{M}^1_{1+E^k}}\right)$$

for any $k \geq 0$ under the condition (C3) with a constant $\kappa(C,t,k,e_0,e_1)$.

Proof. One has

$$\mathbf{E}\left(g_2,\left(\frac{Z_t^h(Z_0^h)-\mu_t(\mu_0)}{\sqrt{h}}\right)^{\otimes 2}\right)=\mathbf{E}\left(g_2,\left(\frac{Z_t^h(Z_0^h)-\mu_t(Z_0^h)}{\sqrt{h}}\right)^{\otimes 2}\right)+\left(g_2,\left(\frac{\mu_t(Z_0^h)-\mu_t(\mu_0)}{\sqrt{h}}\right)^{\otimes 2}\right)$$

$$+2\mathbf{E}\left(g_2,\frac{Z_t^h(Z_0^h)-\mu_t(Z_0^h)}{\sqrt{h}}\otimes\frac{\mu_t(Z_0^h)-\mu_t(\mu_0)}{\sqrt{h}}\right) \tag{7.2}$$

The first term can be rewritten as

$$\frac{1}{h}\mathbf{E}\big(g_2,(Z_t^h(Z_0^h))^{\otimes 2}-(\mu_t(Z_0^h))^{\otimes 2}$$
$$+\mu_t(Z_0^h)\otimes(\mu_t(Z_0^h)-Z_t^h(Z_0^h))+(\mu_t(Z_0^h)-Z_t^h(Z_0^h))\otimes\mu_t(Z_0^h)\big).$$

Under the condition (C1) or (C2) this term can be estimated by

$$\kappa(C,r,e_0,e_1)\|g_2\|_{(1+E^k)^{\otimes 2}}(1+(E^{2k+2},Z_0^h))(1+(E^k,Z_0^h)),$$

due to Theorem 2.1 and (5.2). The second term is estimated by

$$\|g_2\|_{(1+E^k)^{\otimes 2}}(1+(E^{k+1},\mu_0+Z_0^h))\left\|\frac{Z_0^h-\mu_0}{\sqrt{h}}\right\|^2_{1+E^k}$$

by (2.3), and the third term by the obvious combination of these two estimates completing the proof for cases (C1) and (C2). The case (C3) is considered analogously. Namely, the first term in the representation (7.2) is again estimated by Theorem 2.1, and to estimate the second term one uses (5.15) instead of (2.3) and the observation that

$$|(g_2,\nu^{\otimes 2})|\leq\sup_{x_1}\left|(1+E^k(x_1))^{-1}\int\frac{\partial g_2}{\partial x_1}(x_1,x_2)\nu(dx_2)\right|\|\nu\|_{\mathcal{M}^1_{1+E^k}}$$

$$\leq\sup_{x_1,x_2}\left|(1+E^k(x_1))^{-1}(1+E^k(x_2))^{-1}\frac{\partial^2 g_2}{\partial x_1\partial x_2}(x_1,x_2)\right|\|\nu\|^2_{\mathcal{M}^1_{1+E^k}}\leq\|g_2\|_{C^{2,sym}_{(1+E^k)^{\otimes 2}}}\|\nu\|^2_{\mathcal{M}^1_{1+E^k}}.$$

Though the estimates of Proposition 7.1 are sufficient to prove Theorem 2.2, in order to prove the semigroup convergence from Theorem 2.4 one needs a slightly more general estimate, which in turn requires a more general form of LLN, than presented in Theorem 2.1. We shall give now these two extensions.

Remark. Let us stress for clarity that $U_{fl}^{h;0,s}((g_2,.)G)$ means the result of the evolution $U_{fl}^{h;0,s}$ applied to the function of Y given by $(g_2,Y^{\otimes 2})G(Y)$.

Proposition 7.2 *The estimates on the r.h.s. of* (2.5) *and* (2.6) *remain valid, if on the l.h.s. on takes a more general expression, namely*

$$\sup_{s\le t} |T_s^h(GFH)(Y) - G(Y_s)F(Y_s)T_s^h H(Y)|,$$

where $F(Y)$ is as in Theorem 2.1 and both G and H are cylindrical functionals of the form (2.13) *with $f \in C^2(\mathbf{R}^d)$ and all ϕ_j, $j = 1, ..., n$, belonging to $C_{1+E^k}(X)$ and $C^{2,0}_{1+E^k}(X)$ respectively in cases (C1)-(C2) and (C3) (with a constant C depending on the corresponding norms of ϕ_j).*

Proof. As

$$T_t^h(GFH)(Y) - G(Y_t)F(Y_t)T_t^h H(Y)$$
$$= T_t^h(GFH)(Y) - (GFH)(Y_t) + (GF)(Y_t)(H(Y_t) - T_t^h H(Y)),$$

it is enough to consider the case without a function H involved. And in this case looking through the proof of Theorem 2.1 above one sees that it generalizes straightforwardly to give the result required.

Proposition 7.3 *The estimates of Proposition 7.1 remain valid if instead of* (7.1) *one takes a more general expression*

$$\sup_{s\le t} \left|\mathbf{E}[(g_2, F_s^h(Z_0^h, \mu_0))G(F_s^h(Z_0^h, \mu_0))]\right| = \sup_{s\le t} |[U_{fl}^{h;0,s}((g_2, .)G)](F_0^h)|, \tag{7.3}$$

where G is as in the previous Proposition (with a constant C again depending on the norms of ϕ_j in the representation of G as a cylindrical function of the form (2.13)*).*

Proof. It is again obtained by a straightforward generalization of the proof of Proposition 7.1 given above using Proposition 7.2 instead of Theorem 2.1.

The main technical ingredient in the proof of the functional CLT (stated as Theorems 2.5-2.6) is given by the following

Proposition 7.4 *Under condition (C3) for any $k > 1/2$*

$$\sup_{s\le t} \mathbf{E}\|F_s^h(Z_0^h, \mu_0)\|^2_{(L^{2,0}_{2,f_k})'} \le \kappa(C, t, k, e_0, e_1)(1 + (E^{2k+3}, Z_0^h + \mu_0))^2(1 + \|F_0^h\|^2_{(L^{2,0}_{2,f_k})'}). \tag{7.4}$$

Proof. The idea is to represent the l.h.s. of (7.4) in the form of the l.h.s. of (7.1) with an appropriate function g_2. Using the notation $\tilde\nu(x) = \int_x^\infty \nu(dy)$ from the introduction for a finite (signed) measure ν on $\mathbf{R}_+$ (and setting $\tilde\nu(x) = 0$ for $x < 0$) one has

$$\mathcal{F}(\tilde\nu) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-ipy}(\int_y^\infty \nu(dx))dy = \frac{1}{\sqrt{2\pi}} \int_0^\infty \nu(dx) \int_0^x e^{-ipy}\, dy,$$

so that for $f_k(x) = 1 + x^k$

$$\mathcal{F}(f_k\tilde\nu) = (1 + (i\frac{\partial}{\partial p})^k)\mathcal{F}(\tilde\nu) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \nu(dx) \int_0^x (1 + y^k)e^{-ipy}\, dy.$$

Applying (A.3) yields

$$\|\nu\|^2_{(L^{2,0}_{2,f_k}(\mathbf{R}_+))'} = \|f_k\tilde{\nu}\|^2_{H^{-1}(\mathbf{R})} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|\int_0^{\infty}\nu(dx)\int_0^x(1+y^k)e^{-ipy}\,dy\right|^2\frac{dp}{1+p^2}$$

$$= \frac{1}{2\pi}\int_0^{\infty}\int_0^{\infty}\theta_k(x,y)\nu(dx)\nu(dy) \tag{7.5}$$

with

$$\theta_k(x,y) = \int_{-\infty}^{\infty}\left(\int_0^x(1+z^k)e^{-ipz}\,dz\int_0^y(1+w^k)e^{ipw}\,dw\right)\frac{dp}{1+p^2}. \tag{7.6}$$

Clearly

$$\theta_k(x,y)|_{x=0} = \theta_k(x,y)|_{y=0} = 0.$$

Moreover

$$\frac{\partial\theta_k}{\partial x}(x,y) = \int_{-\infty}^{\infty}[(1+x^k)e^{-ipx}\int_0^y(1+w^k)e^{ipw}dw]\frac{dp}{1+p^2}$$

so that

$$|\frac{\partial\theta_k}{\partial x}(x,y)| \le c(k)(1+x^k)(1+y^{k+1}),$$

and

$$|\frac{\partial^2\theta_k}{\partial x\partial y}(x,y)| = \left|\int_{-\infty}^{\infty}(1+x^k)e^{-ipx}(1+y^k)e^{ipy}\frac{dp}{1+p^2}\right| \le c(k)(1+x^k)(1+y^k).$$

Since

$$\frac{\partial^2\theta_k}{\partial x^2}(x,y) = \int_{-\infty}^{\infty}[kx^{k-1}e^{-ipx}\int_0^y(1+w^k)e^{ipw}dw]\frac{dp}{1+p^2}$$
$$-\int_{-\infty}^{\infty}[(1+x^k)e^{-ipx}\int_0^y(1+w^k)(ip)e^{ipw}dw]\frac{dp}{1+p^2},$$

and using integration by parts in the second term yields also

$$|\frac{\partial^2\theta_k}{\partial x^2}(x,y)| \le c(k)[(1+x^k)(1+y^k)+(1+x^{k-1})(1+y^{k+1})].$$

Consequently $\theta_k+\bar{\theta}_k \in C^{2,sym}_{(1+E^{k+1})^{\otimes 2}}$. Therefore, using (7.5) for $\nu = F^h_s(Z^h_0,\mu_0)$ implies that in order to estimate the l.h.s. of (7.4) one needs to estimate the l.h.s. of (7.1) with $g_2 = \theta_k$ given by (7.6).

Though a direct application of Proposition 7.1 does not give the result we need, only a slight modification is required. Namely, representing (7.1) in form (7.2) we estimate the first term precisely like in the proof of Proposition 7.1 and the second term that now equals

$$\left\|\frac{\mu_t(Z^h_0)-\mu_t(\mu_0)}{\sqrt{h}}\right\|^2_{(L^{2,0}_{2,f_k})'}$$

can be estimated using Proposition 5.8 by

$$\kappa(c,t,k,e_0,e_1)(1+E^{k+1},\mu_0+Z^h_0)\|\frac{Z^h_0-\mu_0}{\sqrt{h}}\|^2_{(L^{2,0}_{2,f_k})'}.$$

Estimating the third term in (7.2) again by the combination of the estimates of the first two terms yields (7.4).

8 CLT: Proof of Theorems 2.2 - 2.6

Proof of Theorem 2.2. Recall that we denoted by $U_{fl}^{h;t,r}$ the backward propagator corresponding to the process $F_t^h = (Z_t^h - \mu_t)/\sqrt{h}$. By (3.9), the l.h.s. of (2.11) can be written as

$$\sup_{s\le t}\left|(U_{fl}^{h;0,s}(g,.))(F_0^h) - (U^{0,s}g, F_0^h)\right| = \sup_{s\le t}\left|\int_0^s [U_{fl}^{h;0,\tau}(\Lambda_\tau^h - \Lambda_\tau)U^{\tau,s}(g,.)]\, d\tau\, (F_0^h)\right|.$$

As by (3.10)

$$(\Lambda_\tau^h - \Lambda_\tau)(U^{\tau,s}g,\cdot)(Y) = \frac{\sqrt{h}}{2}\int\int\int (U^{\tau,s}g(y) - U^{\tau,s}g(z_1) - U^{\tau,s}g(z_2))K(z_1,z_2;dy)Y(dz_1)Y(dz_2)$$
$$-\frac{\sqrt{h}}{2}\int\int (U^{\tau,s}g(y) - 2U^{\tau,s}g(z))K(z,z;dy)(\mu_t + \sqrt{h}Y)(dz)$$

(note that the terms with the second and third variational derivatives in (3.10) vanish here, as we apply it to a linear function), the required estimate follows from Proposition 7.1.

Proof of Theorem 2.3. Substituting the function $\Phi_{f_t}^{\phi_1^t,...,\phi_n^t}$ of form (2.13) (with two times continuously differentiable f_t) with a given initial condition $\Phi_r(Y) = \Phi_{f_r}^{\phi_1^r,...,\phi_n^r}(Y)$ at $t=r$ in the equation $\dot{F}_t = -\Lambda_t F_t$ yields

$$\frac{\partial f_t}{\partial t} + \frac{\partial f_t}{\partial x_1}(\dot{\phi}_1^t, Y) + ... + \frac{\partial f_t}{\partial x_n}(\dot{\phi}_n^t, Y)$$
$$= -\frac{1}{2}\int\int\int \sum_{j=1}^n \frac{\partial f_t}{\partial x_j}(\phi_j^t, \delta_y - \delta_{z_1} - \delta_{z_2})K(z_1,z_2;dy)(Y(dz_1)\mu_t(dz_2) + \mu_t(dz_1)Y(dz_2))$$
$$-\frac{1}{4}\int\int\int \sum_{j,l=1}^n \frac{\partial^2 f_t}{\partial x_j \partial x_l}(\phi_j^t\otimes\phi_l^t, (\delta_y - \delta_{z_1} - \delta_{z_2})^{\otimes 2})K(z_1,z_2;dy)\mu_t(dz_1)\mu_t(dz_2)$$

with $f_t(x_1,...,x_n)$ and all its derivatives evaluated at the points $x_j = (\phi_j^t, Y)$ (here and in what follows we denote by dot the derivative d/dt with respect to time). This equation is clearly satisfied whenever

$$\dot{f}_t(x_1,...,x_n) = -\sum_{j,k=1}^n \Pi(t,\phi_j^t,\phi_k^t)\frac{\partial^2 f_t}{\partial x_j \partial x_k}(x_1,...,x_n) \tag{8.1}$$

and

$$\dot{\phi}_j^t(z) = -\int\int (\phi_j^t(y) - \phi_j^t(z) - \phi_j^t(w))K(z,w;dy)\mu_t(dw) = -\Lambda_t\phi_j^t(z)$$

with Π given by (2.19). Consequently

$$OU^{t,r}\Phi_r(Y) = \Phi_t(Y) = (\mathcal{U}^{t,r}f_r)((U^{t,r}\phi_1^r, Y), ..., (U^{t,r}\phi_n^r, Y)), \tag{8.2}$$

where $\mathcal{U}^{t,r}f_r = \mathcal{U}_\Pi^{t,r}f_r$ is defined as the resolving operator to the (inverse time) Cauchy problem of equation (8.1) (it is well defined as (8.1) is just a spatially invariant second order evolution), the resolving operator $U^{t,r}$ is constructed in Sections 4,5, and

$$\Pi(t,\phi_j^t,\phi_k^t) = \Pi(t, U^{t,r}\phi_j^r, U^{t,r}\phi_k^r).$$

All statements of Theorem 2.3 follows from the explicit formula (8.2), the semigroup property of the solution to finite-dimensional equation (8.1) and Propositions 5.4, 5.7.

Proof of Theorem 2.4. The first statement is obtained by a straightforward modification of our proof of Theorem 2.2 above, where one has to use Proposition 7.3 instead of its particular case Proposition 7.1 and to note that all terms in formula (3.10) (that unlike the linear case now become relevant) depend at most quadratically on Y, because for a function Φ of form (2.13)

$$\delta\Phi(Y;x) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\phi_j(x),$$

$$\delta^2\Phi(Y;x,y) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}\phi_j(x)\phi_i(y), \tag{8.3}$$

where the derivatives of f are evaluated at the points $x_j = (\phi_j^t, Y)$.

The second statement follows by the usual approximation of a general Φ by those given by (2.13) with smooth f.

Proof of Theorem 2.5. The characteristic function of Φ_t^h is

$$g^h_{t_1,...,t_n}(p_1,...,p_n) = \mathbf{E}\exp\{\sum_{j=1}^{n}(\phi_j, F^h_{t_j}(Z^h_0,\mu_0))\} = U_{fl}^{h;0,t_1}\Phi_1 ... U_{fl}^{h;t_{n-2},t_{n-1}}\Phi_{n-1}U_{fl}^{h;t_{n-1},t_n}\Phi_n(F_0^h),$$

where $\Phi_j(Y) = \exp\{ip_j(\phi_j, Y)\}$. Let us show that it converges to the characteristic function

$$g_{t_1,...,t_n}(p_1,...,p_n) = OU^{0,t_1}\Phi_1 ... OU^{t_{n-2},t_{n-1}}\Phi_{n-1}OU^{t_{n-1},t_n}\Phi_n(F_0) \tag{8.4}$$

of a Gaussian random variable. For $n = 1$ it follows from Theorem 2.4. For $n > 1$ one can write

$$g^h_{t_1,...,t_n}(p_1,...,p_n) - g_{t_1,...,t_n}(p_1,...,p_n)$$

$$= \sum_{j=1}^{n} U_{fl}^{h;0,t_1}\Phi_1 ... U_{fl}^{h;t_{j-2},t_{j-1}}\Phi_{j-1}(U_{fl}^{h;t_{j-1},t_j} - OU^{t_{j-1},t_j})\Phi_j OU^{t_j,t_{j+1}} ... OU^{t_{n-1}t_n}\Phi_n. \tag{8.5}$$

By Theorem 2.4 we know that for any $j = 1,...,n$

$$\Psi_j^{p_j,...,p_n}(Y) = (U_{fl}^{h;t_{j-1},t_j} - OU^{t_{j-1},t_j})\Phi_j OU^{t_j,t_{j+1}} ... OU^{t_{n-1}t_n}\Phi_n(Y)$$

converge to zero as $\sqrt{h}$ as $h \to 0$ uniformly on Y from bounded domains of $\mathcal{M}^1_{1+E^k}$. We have to show that

$$U_{fl}^{h;t_{j-2},t_{j-1}}\Phi_{j-1}\Psi_j(Y) = \mathbf{E}^h_Y(\Phi_{j-1}\Psi_j(Y_{t_{j-1}})) \tag{8.6}$$

tends to zero, where $\mathbf{E}^h_Y$ is of course the expectation with respect to the fluctuation process started in Y at time t_{j-2}. The last expression can be written as

$$\mathbf{E}^h_Y((\mathbf{1}_{\{\|Y_{t_{j-1}}\|_{(L^{2,0}_{2,f_{k+2}})'} \le K\}}\Phi_{j-1}\Psi_j)(Y_{t_{j-1}})) + \mathbf{E}^h_Y((\mathbf{1}_{\{\|Y_{t_{j-1}}\|_{(L^{2,0}_{2,f_{k+2}})'} > K\}}\Phi_{j-1}\Psi_j)(Y_{t_{j-1}})). \tag{8.7}$$

For Y from a bounded subset of $(L^{2,0}_{2,f_{k+2}})'$ the second term can be made arbitrary small by choosing large enough K due to Proposition 7.4. Due to the natural continuous inclusion

$C^{m,0}_{f_k} \subset L^{m,0}_{2,f_{k+\alpha}}$, $m = 1,2$, $\alpha > 1/2$ one gets by duality a continuous projection $(L^{2,0}_{2,f_k})' \mapsto \mathcal{M}^2_{f_{k-\alpha}} \subset \mathcal{M}^1_{f_{k-\alpha}}$ for $k > 1/2$, $\alpha \in (1/2, k)$. Hence a bounded set in $(L^{2,0}_{2,f_{k+2}})'$ is also bounded in $\mathcal{M}^1_{f_{k+2-\alpha}}$, so that there $\Phi_{j-1}\Psi_j(Y_{t_{j-1}})$ is small of order $\sqrt{h}$, implying that the first term in (8.7) is small. Consequently expression (8.6) tends to zero uniformly for Y from bounded domain of $(L^{2,0}_{2,f_{k+2}})'$, $k > 1/2$. This implies that all terms in (8.5) tend to zero as $h \to 0$.

It remains to check that (8.4) is given by (2.18), which is done by induction in n using Theorem 2.3 and an obvious explicit formula

$$\mathcal{U}^{t,r} f(x) = \exp\{i \sum_{j=1}^n p_j x_j - \sum_{j,k=1}^n p_j p_k \int_t^r \Pi(s, \phi^s_j, \phi^s_k)\, ds\}$$

for the solution of the Cauchy problem of the diffusion equation (8.1) with $f(x) = \exp\{i \sum_{j=1}^n p_j x_j\}$. For instance,

$$(OU^{t_{n-1},t_n}\Phi_n)(Y) = (\mathcal{U}^{t_{n-1},t_n} f_n)(U^{t_{n-1},t_n}\phi_n, Y)$$

$$= \exp\{ip_n(U^{t_{n-1},t_n}\phi_n, Y) - p_n^2 \int_{t_{n-1}}^{t_n} \Pi(s, U^{s,t_n}\phi_n, U^{s,t_n}\phi_n)\, ds\}$$

where $f_n(x) = \exp\{ip_n x\}$, and hence

$$OU^{t_{n-2},t_{n-1}}(\Phi_{n-2} OU^{t_{n-1},t_n}\Phi_n)(Y)$$

$$= \exp\{i(p_{n-1}U^{t_{n-2},t_{n-1}}\phi_{n-1} + p_n U^{t_{n-2},t_n}\phi_n, Y) - p_n^2 \int_{t_{n-2}}^{t_n} \Pi(s, U^{s,t_n}\phi_n, U^{s,t_n}\phi_n)\, ds\}$$

$$\times \exp\{-\int_{t_{n-2}}^{t_{n-1}} [p^2_{n-1}\Pi(s, U^{s,t_{n-1}}\phi_{n-1}, U^{s,t_{n-1}}\phi_{n-1}) + 2p_{n-1}p_n\Pi(s, U^{s,t_{n-1}}\phi_{n-1}, U^{s,t_n}\phi_n)]\, ds\}.$$

The proof is complete.

Proof of Theorem 2.6.

(i) Notice first that applying Dynkin's formula to the Markov process Z^h_t one finds that for a $\phi \in C_{1+E^k}(X)$

$$M^h_\phi(t) = (\phi, Z^h_t) - (\phi, Z^h_0) - \int_0^t (L_h(\phi, .))(Z^h_s) ds$$

is a martingale, since all three terms here are integrable, due to formula (3.7) and the assumption $Z^h_0 \in \mathcal{M}_{1+E^{k+5}}$. Hence (ϕ, F^h_t) is a semimartingale and

$$(\phi, F^h_t) = \frac{M^h_\phi}{\sqrt{h}} + V^h_\phi(t)$$

with

$$V^h_\phi(t) = \frac{1}{\sqrt{h}}[(\phi, Z^h_0) + \int_0^t (L_h(\phi, .))(Z^h_s) ds - (\phi, \mu_t)]$$

is the canonical representation of the semimartingale (ϕ, F^h_t) into the sum of a martingale and a predictable process of bounded variation that is also continuous and integrable. (It implies, in particular that (ϕ, F^h_t) belongs to the class of special semimartingales.)

As we know already the convergence of finite dimensional distributions, to prove (i) one has to show that the distribution on the Skorohod space of càdlàg functions of the semimartingale (ϕ, F_t^h) is tight, which according to Aldous-Rebolledo Criterion (see e.g. [9], [36], we cite the formulation from [9]) amounts to showing that given a sequence of $h_n \to 0$ as $n \to \infty$ and a sequence of stopping times τ_n bounded by a constant T and an arbitrary $\epsilon > 0$ there exist $\delta > 0$ and $n_0 > 0$ such that

$$\sup_{n \geq n_0} \sup_{\theta \in [0,\delta]} P\left[|V^{(n)}(\tau_n + \theta) - V^{(n)}(\tau_n)| > \epsilon\right] \leq \epsilon,$$

and

$$\sup_{n \geq n_0} \sup_{\theta \in [0,\delta]} P\left[|Q^{(n)}(\tau_n + \theta) - Q^{(n)}(\tau_n)| > \epsilon\right] \leq \epsilon,$$

where $V^n(t)$ is a shorter notation for $V_\phi^{h_n}$ and $Q^n(t)$ is the quadratic variation of the martingale $M_\phi^{h_n}(t)$. Notice that it is enough to show the tightness of (ϕ, F_t^h) for a dense subspace of the test functions ϕ. Thus we can and will consider now only the bounded ϕ.

To get a required estimate for $V^n(t)$ observe that by (3.7)

$$\frac{d}{dt}V^n(t) = \frac{1}{\sqrt{h}}[(L_h(\phi,.))(Z_s^h) - (\phi, \dot{\mu}_t)] = -\frac{\sqrt{h}}{2}\int\int[\phi(y) - 2\phi(z)]K(z,z;dy)Z_t^h(dz)$$

$$+\frac{1}{2\sqrt{h}}\int\int\int[\phi(y) - \phi(z_1) - \phi(z_2)]K(z_1,z_2;dy)[Z_t^h(dz_1)Z_t^h(dz_2) - \mu_t(dz_1)\mu_t(dz_2)].$$

The first term here is clearly uniformly bounded for $h \to 0$, and the second term can be written as

$$\frac{1}{2}\int\int\int[\phi(y) - \phi(z_1) - \phi(z_2)]K(z_1,z_2;dy)[F_t^h(dz_1)Z_t^h(dz_2) + \mu_t(dz_1)F_t^h(dz_2)]. \tag{8.8}$$

Applying Doob's maximal inequality to the martingale

$$(\phi, F_t^h) - (\phi, F_0^h) - \int_0^t (\Lambda_s^h(\phi,.))(Z_s^h)ds$$

in combination with Proposition 7.4 shows that (8.8) can be made bounded with an arbitrary small probability, implying the required estimate for $V^n(t)$.

Let us estimate the quadratic variation by the same arguments as in [18]. Namely, as the process (ϕ, F_t^h) for each h is the sum of a differentiable process and a pure jump process, both having locally finite variation, its quadratic variation coincides with that of $M_\phi^{h_n}(t)$ and is known to equal the sum of the squares of the sizes of all its jumps (see e.g. Theorem 26.6 in [14]), so that

$$Q^{(n)}(t) - Q^{(n)}(\tau) = \sum_{s \in [\tau,t]} (\phi, F_s^{h_n} - F_{s-}^{h_n})^2 = \frac{1}{h}\sum_{s \in [\tau,t]} (\phi, Z_s^{h_n} - Z_{s-}^{h_n})^2.$$

As each jump of Z_s^h is the change of $h\delta_x + h\delta_y$ to $h\delta_{x+y}$ one concludes that

$$|Q^{(n)}(t) - Q^{(n)}(\tau)| \leq h \sup|\phi||N_t - N\tau|$$

with N_t denoting the number of jumps on the interval $[0, t]$. By the Lévy formula for Markov chains (see e.g. [4]) the process $N_t - \int_0^t a(Z_s^h)\,ds$ is a martingale, where $a(Y)$ denotes the intensity of jumps at Y, given by (2.4). Hence, using the optional sampling theorem and (2.4) implies that

$$\mathbf{E}(N_t - N_\tau) = \mathbf{E}\int_\tau^t a(Z_s^h)\,ds \le 3Ch^{-1}e_0(e_1+e_0)\mathbf{E}(t-\tau),$$

and consequently

$$\mathbf{E}|Q^n(t) - Q^n(\tau)| \le 3C\|\phi\|e_0(e_1+e_0)\theta$$

uniformly for all $< t-\theta < \tau < t$. Hence by Chebyshev inequality the required estimate for Q^n follows.

(ii) By Theorem 2.5 the limiting process is uniquely defined whenever it exists. Hence one only needs to prove the tightness of the family of normalized fluctuations F_t^h. Again due to the existence of finite dimensional limits and general convergence theorems, to prove tightness it is enough to establish the following compact containment condition (see either a result of [31] specially designed to show convergence in Hilbert spaces, or a more general result on convergence of complete separable metric space valued processes in [11] or [9], see also [10]): for every $\epsilon > 0$ and $T > 0$ there exists $K > 0$ such that for any h

$$P(\sup_{t\in[0,T]} \|F_t^h\|_{(L^{2,0}_{2,f_k})'} > K) \le \epsilon.$$

To this end, let us introduce a regularized square root function R, i.e. $R(x)$ is an infinitely smooth increasing function $\mathbf{R}_+ \mapsto \mathbf{R}_+$ such that $R(x) = \sqrt{x}$ for $x > 1$, and the corresponding "regularized norm" functional on $(L^{2,0}_{2,f_k})'$:

$$G(Y) = R((Y,Y)_{(L^{2,0}_{2,f_k})'}) = R((\theta_k, Y\otimes Y)),$$

where θ_k is given by (7.5) (see Proposition 7.4). By Dynkin's formula one can conclude that the process

$$M_t = G(F_t^h) - G(F_0^h) - \int_0^t \Lambda_s^h G(F_s^h)\,ds$$

is a martingale whenever all terms in this expression have finite expectations. (Note that we use here a more general than usual version of Dynkin's formula with a time dependent generator; the reduction of time nonhomogeneous case to the standard situation by including time as an additional coordinate of a Markov process under consideration is explained e.g. in [12].) Expectation of $G(F_t^h)$ is bounded by Proposition 7.4. Moreover, taking into account (8.3) and the fact that $R^{(k)}(s) = O(s^{(1/2)-k})$ for $s \ge 1$, one sees from formulas (3.10) and (2.9) that $\Lambda_s^h G(F_s^h)$ grows at most quadratically in F_s^h, which again by Proposition 7.4 implies the uniform boundedness of the expectation of this term. Applying to M_t Doob's maximal inequality yields the required compact containment completing the proof of the theorem.

A Notations for weighted spaces of functions and distributions

For a positive measurable function f on a topological space T we denote by $C_f = C_f(T)$ and $B_f = B_f(T)$ (omitting T when no ambiguity may arise) the Banach spaces of continuous and measurable functions on T respectively having finite norm

$$\|\phi\|_f = \|\phi\|_{C_f(T)} = \sup_x (|\phi(x)|/f(x)).$$

By $C_{f,\infty} = C_{f,\infty}(T)$ and $B_{f,\infty} = B_{f,\infty}(T)$ we denote the subspaces of C_f and B_f respectively consisting of functions ϕ such that $(\phi/f)(x) \to 0$ as $f(x) \to \infty$. If f is a continuous function on a locally compact space X such that $f(x) \to \infty$, as $x \to \infty$, then the dual space to $C_{f,\infty}(X)$ is given by the space $\mathcal{M}_f(X)$ of Radon measures on X with the norm $\|Y\|_f = \sup\{(\phi, Y) : \|\phi\|_f \leq 1\}$.

We shall need also the weighted L_p spaces. Namely, define $L_{p,f} = L_{p,f}(T)$ as the space of measurable functions g on a measurable space T having finite norm $\|g\|_{L_{p,f}} = \|g/f\|_{L_p}$.

For $X = \mathbf{R}_+ = \{x > 0\}$ we shall use also smooth functions. For a positive f we denote by $C_f^{1,0} = C_f^{1,0}(X)$ the Banach space of continuously differentiable functions ϕ on $X = \mathbf{R}_+$ such that $\lim_{x\to 0} \phi(x) = 0$ and the norm

$$\|\phi\|_{C_f^{1,0}(X)} = \|\phi'\|_{C_f(X)}$$

is finite. By $C_f^{2,0} = C_f^{2,0}(X)$ we denote the space of two-times continuously differentiable functions such that $\lim_{x\to 0} \phi(x) = 0$ and the norm

$$\|\phi\|_{C_f^{2,0}(X)} = \|\phi'\|_f + \|\phi''\|_f$$

is finite. By $\mathcal{M}_f^1(X)$ and $\mathcal{M}_f^2(X)$ we shall denote the Banach dual spaces to $C_f^{1,0}$ and $C_f^{2,0}$ respectively. Actually we need only the topology they induce on (signed) measures so that for $\nu \in \mathcal{M}(X) \cap \mathcal{M}_f^i(X)$, $i = 1, 2$,

$$\|\nu\|_{\mathcal{M}_f^i(X)} = \sup\{(\phi, \nu) : \|\phi\|_{C_f^{i,0}(X)} \leq 1\}.$$

Similarly one defines the spaces $L_{p,f}^{1,0}$ and $L_{p,f}^{2,0}$, $p \geq 1$, as the spaces of absolutely continuous functions ϕ on $X = \mathbf{R}_+$ such that $\lim_{x\to 0} \phi(x) = 0$ with the norms respectively

$$\|\phi\|_{L_{p,f}^{1,0}(X)} = \|\phi'\|_{L_{p,f}(X)} = \|\phi'/f\|_{L_p(X)}, \quad \|\phi\|_{L_{p,f}^{2,0}(X)} = \|\phi'/f\|_{L_p(X)} + \|(\phi'/f)'\|_{L_p(X)},$$

as well as their dual $(L_{p,f}^{1,0})'$ and $(L_{p,f}^{2,0})'$.

As an important example let us estimate two of these norms for the Dirac measure δ_x on $\mathbf{R}_+$, $x > 0$ and the function $f(y) = f_k(y) = 1 + y^k$:

$$\|\delta_x\|_{\mathcal{M}_{f_k}^1(\mathbf{R}_+)} = \sup\{\int_0^x g(y)\,dy : \|g\|_{C_{f_k}} \leq 1\} = x + x^{k+1}/(k+1);$$

$$\|\delta_x\|_{(L_{2,f_k}^{1,0})'(\mathbf{R}_+)} = \sup\{\int_0^x g(y)\,dy : \|g/f_k\|_{L_2} \leq 1\} \leq \sqrt{\int_0^x f_k^2(y)\,dy} \leq c(k)\sqrt{x} f_k(x). \quad \text{(A.1)}$$

Not every $\nu \in \mathcal{M}(X)$ belongs to $\mathcal{M}_f^1(X)$ or $\mathcal{M}_f^2(X)$. Suppose that f is non-decreasing and $\nu \in \mathcal{M}(X)$ is such that

$$\tilde{\nu}(x) = \int_x^\infty \nu(dy) = o(1)(xf(x))^{-1}, \quad x \to \infty. \quad \text{(A.2)}$$

Then by integration by parts for $g \in C_f^{1,0}(\mathbf{R}_+)$

$$(g, \nu) = -\int_0^\infty g(x) d\tilde{\nu}(x) = \int_0^\infty g'(x)\tilde{\nu}(x) dx$$

(the boundary term vanish by (A.2)), so that

$$\|\nu\|_{\mathcal{M}_f^1(X)} = \|\tilde{\nu}\|_{L_{1,1/f}}$$

and

$$\|\nu\|_{\mathcal{M}_f^2(X)} = \sup\{(\phi, \tilde{\nu}) : \|\phi\|_{C_f} + \|\phi'\|_{C_f} \leq 1\}.$$

Similarly, as

$$\int_0^x \phi(s) ds \leq \|\phi\|_{L_{p,f}} \left(\int_0^x f^q(y) dy\right)^{1/q}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

it follows that if $\nu \in \mathcal{M}(X)$ is such that

$$\tilde{\nu} = o(1) \left(\int_0^x f^q(y) dy\right)^{-1/q}, \quad x \to \infty,$$

then

$$\|\nu\|_{(L_{p,f}^{1,0})'} = \|\tilde{\nu}\|_{L_{q,1/f}}, \quad \frac{1}{p} + \frac{1}{q} = 1,$$

$$\|\nu\|_{(L_{p,f}^{2,0})'} = \sup\{(\psi, \tilde{\nu}) : \|\psi/f\|_{L_p} + \|(\psi/f)'\|_{L_p} \leq 1\}, \quad p > 1.$$

In particular, recalling that the usual Sobolev Hilbert spaces $H^k(\mathbf{R})$ are defined as the completion of the Schwarz space $S(\mathbf{R})$ with respect to the scalar product

$$(f, g)_{H^k} = (f, (1 - \Delta)^k g)_{L_2} = (\mathcal{F}f, (1 + p^2)^k \mathcal{F}g)_{L_2},$$

where

$$(\mathcal{F}(f))(p) = (2\pi)^{-1/2} \int_{\mathbf{R}} e^{-ipx} f(x)\, dx$$

denotes the usual Fourier transform, and that by duality $(H^k)' = H^{-k}$ it follows that

$$\|\nu\|_{(L_{2,f}^{2,0})'} = \sup\{(\psi, \tilde{\nu}) : \|\psi/f\|_{H^1} \leq 1\} = \sup\{(\phi, f\tilde{\nu}) : \|\phi\|_{H^1} \leq 1\}$$

$$= \|f\tilde{\nu}\|_{H^{-1}} = \sqrt{\int_{-\infty}^\infty |\mathcal{F}(f\tilde{\nu})(p)|^2 \frac{dp}{1 + p^2}}. \tag{A.3}$$

This formula is used in Section 7.

It is useful to observe that by the Sobolev embedding lemma one has the inclusion $L_2^{2,0} \subset C^{1,0}$ and hence also $L_{2,f_k}^{2,0} \subset C_{f_k}^{1,0}$ for arbitrary $k > 0$ implying by duality the inclusion $\mathcal{M}_{f_k}^1 \subset (L_{2,f_k}^{2,0})'$.

By $C^{sym}(X^k)$ we denote the Banach space of symmetric (with respect to all permutations of its arguments) continuous bounded functions on X^k, and by $C^{sym}(\mathcal{X})$- the Banach space of continuous bounded functions on $\mathcal{X}$ whose restrictions on each X^k belong to $C^{sym}(X^k)$. For a function f on X we denote by $f^\otimes$ its natural lifting on $\mathcal{X}$, i.e. $f^\otimes(x_1, \ldots, x_n) = f(x_1) \cdots f(x_n)$.

If f is a positive function on $X^m = \mathbf{R}_+^m$, we denote by $C_f^{1,sym}(X^m)$ (respectively $C_f^{2,sym}(X^m)$) the space of symmetric continuous differentiable functions g on X^m (respectively two-times continuously differentiable) vanishing whenever at least one argument vanishes, with the norm

$$\|g\|_{C_f^{1,sym}(X^m)} = \left\| \frac{\partial g}{\partial x_1} \right\|_{C_f(X^m)} = \sup_{x,j} \left(\left| \frac{\partial g}{\partial x_j} \right| (f^{-1}) \right) (x)$$

and respectively

$$\|g\|_{C_f^{2,sym}(X^m)} = \left\| \frac{\partial g}{\partial x_1} \right\|_{C_f(X^m)} + \left\| \frac{\partial^2 g}{\partial x_1^2} \right\|_{C_f(X^m)} + \left\| \frac{\partial^2 g}{\partial x_1 \partial x_2} \right\|_{C_f(X^m)}.$$

B Three lemmas

We present here three general (not connected to each other) analytic facts used in the main body of the paper. Recall that classes $C^1(\mathcal{M}_f(X))$ were defined in the introduction.

Lemma B.1 *(i) If $F \in C^1(\mathcal{M}_f(X))$ and $Y, \xi \in \mathcal{M}_f(X)$, then*

$$F(Y+\xi) - F(Y) = \int_0^1 (\delta F(Y + s\xi; \cdot), \xi)\, ds. \tag{B.1}$$

(ii) If $F \in C^2(\mathcal{M}_f(X), \phi)$ or $F \in C^3(\mathcal{M}_f(X), \phi)$, the following Taylor expansion holds respectively:

$$(a) \qquad F(Y+\xi) - F(Y) = (\delta F(Y; \cdot), \xi) + \int_0^1 (1-s)(\delta^2 F(Y + s\xi; \cdot, \cdot), \xi \otimes \xi)\, ds,$$

$$\begin{aligned}(b) \qquad F(Y+\xi) - F(Y) &= (\delta F(Y; \cdot), \xi) + \frac{1}{2}(\delta^2 F(Y; \cdot, \cdot), \xi \otimes \xi) \\ &+ \frac{1}{2} \int_0^1 (1-s)^2 (\delta^3 F(Y + s\xi; \cdot, \cdot, \cdot), \xi^{\otimes 3})\, ds.\end{aligned} \tag{B.2}$$

(iii) Let $\phi \le f$. If $t \mapsto \mu_t \in \mathcal{M}_f(X)$ is continuous in the $*$-weak topology of $\mathcal{M}_f(X)$ and is continuously differentiable in the $*$-weak topology of $\mathcal{M}_\phi(X)$, then for any $F \in C^1(\mathcal{M}_f(X), \phi)$

$$\frac{d}{dt} F(\mu_t) = (\delta F(\mu_t; \cdot), \dot{\mu}_t). \tag{B.3}$$

Proof. (i) Using the representation

$$F(Y + s(\delta_x + \delta_y)) - F(Y) = F(Y + s\delta_x) - F(Y) + \int_0^s \delta F(Y + s\delta_x + p\delta_y; y)\, dp$$

for arbitrary points x, y and the uniform continuity of $\delta F(Y + s\delta_x + p\delta_y; y)$ in s, p allows to deduce from (1.9) the existence of the limit

$$\lim_{s \to 0_+} \frac{1}{s}(F(Y + s(\delta_x + \delta_y)) - F(Y)) = \delta F(Y; x) + \delta F(Y; y).$$

Extending similarly to the arbitrary number of points one obtains (B.1) for ξ being an arbitrary finite sum of the Dirac measures $\delta_{x_1} + ... + \delta_{x_n}$.

Assume now that $\xi \in \mathcal{M}_f(X)$ and $\xi_k \to \xi$ as $k \to \infty$ $\star$-weakly in $\mathcal{M}_f(X)$, where all ξ_k are finite sums of Dirac measures. We are going to pass to the limit $k \to \infty$ in the equation (B.1) written for ξ_k. As $F \in C(\mathcal{M}_f)$ one has

$$F(Y + \xi_k) - F(Y) \to F(Y + \xi) - F(Y), \quad k \to \infty.$$

Next, the difference

$$\int_0^1 (\delta F(Y + s\xi_k; \cdot), \xi_k) \, ds - \int_0^1 (\delta F(Y + s\xi; \cdot), \xi) \, ds$$

can be written as

$$\int_0^1 (\delta F(Y + s\xi_k; \cdot), \xi_k - \xi) \, ds + \int_0^1 (\delta F(Y + s\xi_k; \cdot) - \delta F(Y + s\xi; \cdot), \xi) \, ds.$$

The second term tends to zero, because by our assumption the variational derivation δF maps $\mathcal{M}_f(X)$ continuously to $C_{f,\infty}(X)$. The first term tends to zero, because $\xi_k \to \xi$ weakly and the family of functions $\delta F(Y + s\xi_k; .)$ is compact in $C_{f,\infty}(X)$ (which is again due to the assumed continuity of the derivation δF).

Statement (ii) is straightforward from the usual Taylor expansion. Turning to (iii) observe that

$$\frac{d}{dt} F(\mu_t) = \lim_{h \to 0} \frac{1}{h} (F(\mu_{t+h}) - F(\mu_t)),$$

which by (i) and the assumed continuous differentiability can be written as

$$\lim_{h \to 0} \int_0^1 \left(\delta F(\mu_t + s(\mu_{t+h} - \mu_t); .), \frac{1}{h} \int_0^h \dot{\mu}_{t+\tau} d\tau \right) ds.$$

We want to show that it equals the r.h.s. of (B.3). We have

$$\int_0^1 \left(\delta F(\mu_t + s(\mu_{t+h} - \mu_t); .), \frac{1}{h} \int_0^h \dot{\mu}_{t+\tau} d\tau \right) ds - (\delta F(\mu_t; \cdot), \dot{\mu}_t)$$

$$= \left(\delta F(\mu_t; .), \frac{1}{h} \int_0^h \dot{\mu}_{t+\tau} d\tau - \dot{\mu}_t \right) + \int_0^1 \left(\delta F(\mu_t + s(\mu_{t+h} - \mu_t); .) - \delta F(\mu_t; \cdot), \frac{1}{h} \int_0^h \dot{\mu}_{t+\tau} d\tau \right) ds.$$

The first term here tends to zero as $h \to 0$ by the weak continuity of $\dot{\mu}_t$. The second term tends to zero, because the family of measures $h^{-1} \int_0^h \dot{\mu}_{t+\tau} d\tau$ is bounded and hence compact in the $\star$-weak topology of $\mathcal{M}_\phi(X)$.

Lemma B.2 *Suppose S is a compact subset of a linear topological space Y (we are interested in the case when Y is a topological dual to a Banach space equipped with its $\star$-weak topology) and Z_t is a Markov process on S specified by its Feller semigroup Ψ_t on $C(S)$ with a bounded generator A. Let $\Omega_t(z) = (z - \xi_t)/a$ be a family of linear transformation on Y, where a is a positive constant and ξ_t, $t \geq 0$, is a differentiable curve in Y. Let*

$$\Omega_{[0,T]}(S) = \cup_{t \in [0,T]} \Omega_t(S)$$

for $T>0$. Then $Y_t=\Omega_t(Z_t)$, $t\in[0,T]$, is a Markov process in $\Omega_{[0,T]}(S)$ for any $T>0$ with the dynamics of averages (propagator)

$$U^{s,t}f(y)=\mathbf{E}_{s,y}f(Y_t)$$

given by the formula

$$U^{s,t}f(y)=\Omega_s^{-1}\Psi_{t-s}\Omega_t f(y) \tag{B.4}$$

for any $f\in C(\Omega_{[0,T]}(S))$, $t\leq T$, where $\Omega_t f(y)=f(\Omega_t(y))$. Moreover, if such a function f is uniformly continuously differentiable in the direction $\dot{\xi}_t$, i.e. if the limit

$$\lim_{\tau\to 0}\frac{1}{\tau}\left(f(\Omega_{t+\tau}(y))-f(\Omega_t(y))\right)=-\frac{1}{a}(\nabla_{\dot{\xi}_t}f)(\Omega_t(y))=\nabla_{\dot{\xi}_t}f(\Omega_t(y))$$

exists and is uniform in $\Omega_{[0,T]}(S)$, then for all $s\leq t$

$$\frac{d}{dt}U^{s,t}f=U^{s,t}\Lambda_t f, \tag{B.5}$$

where the operator Λ_t is given by the formula

$$\Lambda_t f=\Omega_t^{-1}A\Omega_t f-\frac{1}{a}\nabla_{\dot{\xi}_t}f. \tag{B.6}$$

Proof. Formula (B.4) follows from the definitions of Ψ_t and Ω_t. Formulas (B.5), (B.7) follow by differentiating (B.4) using the product rule and taking into account that the derivative ∇f is supposed to be uniform.

Remark. Similarly, using the identity

$$\Omega_t^{-1}\nabla_{\dot{\xi}_t}\Omega_t=a^{-1}\nabla_{\dot{\xi}_t},$$

one shows that

$$\frac{d}{ds}U^{s,t}f=-\Lambda_s U^{s,t}f, \tag{B.7}$$

holds for $s=t$. However, to extend this to $s<t$ one needs some additional assumptions on the smoothness of the semigroup Ψ_t.

Lemma B.3 *[18] Let Y be a measurable space and the mapping $t\mapsto\mu_t$ from $[0,T]$ to $\mathcal{M}(Y)$ is continuously differentiable in the sense of the norm in $\mathcal{M}(Y)$ with a (continuous) derivative $\dot{\mu}_t=\nu_t$. Let σ_t denote a density of μ_t with respect to its total variation $|\mu_t|$, i.e. the class of measurable functions (equivalence is defined as the a.s. equality with respect to the measure $|\mu_t|$) taking three values $-1,0,1$ and such that $\mu_t=\sigma_t|\mu_t|$ and $|\mu_t|=\sigma_t\mu_t$ almost surely with respect to $|\mu_t|$. Then there exists a measurable function $f_t(x)$ on $[0,T]\times Y$ such that f_t is a representative of class σ_t for any $t\in[0,T]$ and*

$$\|\mu_t\|=\|\mu_0\|+\int_0^t ds\int_Y f_s(y)\nu_s(dy).$$

We refer for a proof to the Appendix of [18] noting only that f_t could be chosen as such a representative of σ_t, which is at the same time a representative of the class of the densities of ν_t^s with respect to its total variation measure $|\nu_t^s|$, where ν_t^s is a singular part of ν_t in its Lebesgue decomposition with respect to $|\mu_t|$.

C On the evolutions with integral generators

Here we present an analytic study of evolutions with integral generators that are obtained as certain perturbations of positivity preserving evolutions. As always, it is assumed that X is a locally compact space (though this assumption is used only in Theorem C.2, other statement being valid for arbitrary topological spaces).

We shall start with the problem

$$\dot{u}_t(x) = A_t u_t(x) = \int u_t(z)\nu_t(x;dz) - a_t(x)u_t(x), \quad u_r(x) = \phi(x), \quad t \geq r \geq 0, \tag{C.1}$$

where ϕ and a_t are given measurable functions on X such that a_t is non-negative and locally bounded in t for each x, $\nu_t(x,\cdot)$ is a given family of finite (non-negative) measures on X depending measurably on $t \geq 0$, $x \in X$, and such that $\sup_{t\in[0,T]} \|\nu_t(x,\cdot)\|$ is bounded for arbitrary T and x.

Clearly equation (C.1) is formally equivalent to the integral equation

$$u_t(x) = I^r_\phi(u)_t = e^{-(\xi_t(x)-\xi_r(x))}\phi(x) + \int_r^t e^{-(\xi_t(x)-\xi_s(x))} L_s u_s(x)ds, \tag{C.2}$$

where $\xi_t(x) = \int_0^t a_s(x)ds$ and $L_t v(x) = \int v(z)\nu_t(x,dz)$.

We shall look for the solutions of (C.2) in the class of functions $u_t(x)$, $t \geq r$, that are continuous in t (for each x), measurable in x and such that the integral in the expression for $L_s u_s$ is well defined in the Lebesgue sense. Basic obvious observation about (C.2) is the following: the iterations of the mapping I^r_ϕ form (C.2) are connected with the partial sums

$$S^{t,r}_m\phi = \left[e^{-(\xi_t-\xi_r)} + \sum_{l=1}^m \int_{r\leq s_l\leq\cdots\leq s_1\leq t} e^{-(\xi_t-\xi_{s_1})} L_{s_1}\cdots L_{s_{l-1}} e^{-(\xi_{s_{l-1}}-\xi_{s_l})} L_{s_l} e^{-(\xi_{s_l}-\xi_r)} ds_1\cdots ds_l \right]\phi$$

(where $e^{-\xi_t}$ designates the operator of multiplication by $e^{-\xi_t(x)}$) of the perturbation series solution $S^{t,r} = \lim_{m\to\infty} S^{t,r}_m$ to (C.2) by

$$(I^r_\phi)^m(\phi)_t = S^{t,r}_{m-1}\phi + \int_{r\leq s_m\leq\cdots\leq s_1\leq t} e^{-(\xi_t-\xi_{s_1})} L_{s_1}\cdots L_{s_{m-1}} e^{-(\xi_{s_{m-1}}-\xi_{s_m})} L_{s_m}\phi ds_1\cdots ds_m. \tag{C.3}$$

Lemma C.1 *Suppose*

$$A_t\psi(x) \leq c\psi(x), \quad t\in[0,T], \tag{C.4}$$

for a strictly positive measurable function ψ on X and a constant $c = c(T)$. Then

$$(I^r_\psi)^m(\psi)_t \leq \left(1 + c(t-r) + \cdots + \frac{1}{m!}c^m(t-r)^m\right)\psi \tag{C.5}$$

for all $0 \leq r \leq t \leq T$, and consequently $S^{t,r}\psi$ is well defined as a convergent series for each t, x and

$$S^{t,r}\psi(x) \leq e^{c(t-r)}\psi(x). \tag{C.6}$$

Proof. This is given by induction in m. Suppose (C.5) holds for m. Since (C.4) implies

$$L_t\psi(x) \le (c + a_t(x))\psi(x) = (c + \dot{\xi}_t(x))\psi(x),$$

it follows that

$$\begin{aligned}(I_\psi^r)^{m+1}(\psi)_t &\le e^{-(\xi_t(x)-\xi_r(x))}\psi(x) \\ &\quad + \int_r^t e^{-(\xi_t(x)-\xi_s(x))}(c+\dot{\xi}_s(x))\left(1 + c(s-r) + \cdots + \frac{1}{m!}c^m(s-r)^m\right)\psi(x)\,ds.\end{aligned}$$

Consequently, as

$$\int_r^t e^{-(\xi_t-\xi_s)}\dot{\xi}_s\frac{1}{l!}(s-r)^l\,ds = \frac{1}{l!}(t-r)^l - \frac{1}{(l-1)!}\int_r^t e^{-(\xi_t-\xi_s)}(s-r)^{l-1}\,ds$$

for $l > 0$, it remains to show that

$$\begin{aligned}&\sum_{l=1}^m c^l\left[\frac{1}{l!}(t-r)^l - \frac{1}{(l-1)!}\int_r^t e^{-(\xi_t-\xi_s)}(s-r)^{l-1}ds\right] + \sum_{l=0}^m c^{l+1}\frac{1}{l!}\int_r^t e^{-(\xi_t-\xi_s)}(s-r)^l ds \\ &\quad \le c(t-r) + \cdots + \frac{1}{(m+1)!}c^{m+1}(t-r)^{m+1}.\end{aligned}$$

But this holds, because the l.h.s. of this inequality equals

$$\sum_{l=1}^m \frac{c^l}{l!}(t-r)^l + \frac{c^{m+1}}{m!}\int_r^t e^{-(\xi_t-\xi_s)}(s-r)^m ds.$$

The following corollary plays an important role in the analysis of Section 5.

Lemma C.2 *Suppose $A_t\psi \le c\psi + \phi$ for positive functions ϕ and ψ and all $t \in [0,T]$. Then*

$$S^{t,r}\psi \le e^{c(t-r)}[\psi + \int_r^t S^{\tau,t}\,d\tau\phi].$$

Proof. Using (C.5) yields

$$\begin{aligned}I_\psi^r(\psi)_t &\le (1 + c(t-r))\psi + \int_r^t e^{-(\xi_t-\xi_s)}\phi\,ds, \\ (I_\psi^r)^2(\psi)_t &\le \left(1 + c(t-r) + \frac{c^2}{2}(t-r)^2\right)\psi + \int_r^t e^{-(\xi_t-\xi_s)}(1 + c(s-r))\phi\,ds \\ &\quad + \int_r^t e^{-(\xi_t-\xi_s)}L_s\int_r^s e^{-(\xi_s-\xi_\tau)}\phi\,d\tau\,ds,\end{aligned}$$

etc, and hence

$$(I_\psi^r)^m(\psi)_t \le e^{c(t-r)}\left[\psi + \int_r^t e^{-(\xi_t-\xi_s)}\phi\,ds + \int_r^t e^{-(\xi_t-\xi_s)}L_s\int_r^s e^{-(\xi_s-\xi_\tau)}\phi\,d\tau ds + \cdots\right]$$

$$= e^{c(t-r)} \left[\psi + \int_r^t d\tau \left(e^{-(\xi_t-\xi_\tau)} + \int_\tau^t e^{-(\xi_t-\xi_s)} L_s e^{-(\xi_s-\xi_\tau)} \, ds + \cdots \right) \phi \right]$$

and the proof is completed by noting that

$$S^{t,r}\psi = \lim_{m\to\infty} S^{t,r}_{m-1}\psi \le \lim_{m\to\infty} (I^r_\psi)^m(\psi)_t.$$

The existence of the solutions to (C.1) and (C.2) can be easily established now.

Proposition C.1 *Under the assumptions of Lemma C.1 the following holds.*

(i) For an arbitrary $\phi \in B_\psi$ the perturbation series $S^{t,r}\phi = \lim_{m\to\infty} S^{t,r}_m \phi$ is absolutely convergent for all t, x, the function $S^{t,r}\phi$ solves (C.2) *and represents its minimal solution (i.e. $S^t\phi \le u$ point-wise for any other solution u to* (C.2)*), and $S^{t,r}\phi(x)$ tends to $S^{\tau,r}\phi(x)$ as $t \to \tau$ uniformly on any set where both a_t and ψ are bounded.*

(ii) The family $S^{t,r}$ form a propagator in $B_\psi(X)$ with the norm

$$\|S^{t,r}\|_\psi \le e^{c(t-r)}. \tag{C.7}$$

Proof. Applying Lemma C.1 separately to the positive and negative part of ϕ one obtains the convergence of series $S^{t,r}\phi$ and the estimate (C.7). Clearly $S^{t,r}\phi$ satisfies (C.2) and it is minimal, as any solution u of this equation satisfies the equation $u_t = (I^r_\phi)^m(u)_t$ and hence (due to (C.3)) also the inequality $u_t \ge S^{t,r}_{m-1}\phi$.

The continuity of $S^{t,r}$ in t follows from the formula

$$S^{t,r}\phi - S^{\tau,r}\phi = (e^{-(\xi_t-\xi_\tau)} - 1)e^{-(\xi_\tau-\xi_r)}\phi$$

$$+ \int_r^\tau (e^{-(\xi_t-\xi_\tau)} - 1) e^{-(\xi_\tau-\xi_s)} L_s S^{s,r}\phi \, ds + \int_\tau^t e^{-(\xi_t-\xi_s)} L_s S^{s,r}\phi \, ds \tag{C.8}$$

for $r \le \tau \le t$.

At last, once the convergence of the series $S^{t,r}$ is proved, the propagator (or Chapman-Kolmogorov) equation (1.10) follows from simple standard manipulations with integrals that we omit.

For the application to time non-homogeneous stochastic processes one needs actually equation (C.1) in inverse time, i.e. the problem

$$\dot u_t(x) = -A_t u_t(x) = -\int u_t(z)\nu_t(x;dz) + a_t(x)u_t(x), \quad u_r(x) = \phi(x), \quad 0 \le t \le r, \tag{C.9}$$

with the corresponding integral equation taking the form

$$u_t(x) = I^r_\phi(u)_t = e^{\xi_t(x)-\xi_r(x)}\phi(x) + \int_t^r e^{\xi_t(x)-\xi_s(x)} L_s u_s(x) ds. \tag{C.10}$$

All the statements of Proposition C.1 (and their proofs) obviously hold for the perturbation series $S^{t,r}$ constructed from (C.10), with the same estimate (C.7), but with the backward propagator equation (1.10) holding for $t \le s \le r$ with S instead of U.

To get a strong continuity of $S^{t,r}$ one usually needs a second bound for A_t. In particular, the following holds.

Proposition C.2 *Suppose now that two measurable functions ψ_1, ψ_2 on X are given both satisfying* (C.4) *and such that (i) $0 < \psi_1 < \psi_2$, (ii) a_t is bounded on any set where ψ_2 is bounded, (iii) $\psi_1 \in B_{\psi_2,\infty}(X)$. Then $S^{t,r}$, $t \leq r$ (constructed above for* (C.9), (C.10)*) is a strongly continuous family of operators in $B_{\psi_2,\infty}(X)$.*

Proof. By Proposition C.1 $S^{t,r}$ are bounded in $B_{\psi_2}(X)$. Moreover, as $S^{t,r}\phi$ tends to ϕ uniformly on the sets where ψ_2 is bounded, it follows that

$$\|S^{t,r}\phi - \phi\|_{\psi_2} \to 0$$

for any $\phi \in B_{\psi_1}(X)$, and hence also for any $\phi \in B_{\psi_2,\infty}(X)$, since $B_{\psi_1}(X)$ is dense in $B_{\psi_2,\infty}(X)$.

Theorem C.1 *Under the assumptions of Proposition C.2 assume additionally that ψ_1, ψ_2 are continuous, a_t is a continuous mapping $t \mapsto C_{\psi_2/\psi_1,\infty}$ and L_t is a continuous mapping from t to bounded operators $C_{\psi_1} \mapsto C_{\psi_2,\infty}$. Then B_{ψ_1} is an invariant core for the propagator $S^{t,r}$ in the sense that*

$$A_r\phi = \lim_{t\to r, t\leq r} \frac{S^{t,r}\phi - \phi}{r-t} = \lim_{s\to r, s\geq r} \frac{S^{r,s}\phi - \phi}{s-r},$$
$$\frac{d}{ds}S^{t,s}\phi = S^{t,s}A_s\phi, \quad \frac{d}{ds}S^{s,r}\phi = -A_sS^{s,r}\phi, \quad t < s < r, \tag{C.11}$$

for all $\phi \in B_{\psi_1}(X)$, with all these limit existing in the Banach topology of $B_{\psi_2,\infty}(X)$. Moreover, C_{ψ_1} and $C_{\psi_2,\infty}$ are invariant under $S^{t,r}$, so that C_{ψ_1} is an invariant core of the strongly continuous propagator $S^{t,r}$ in $C_{\psi_2,\infty}$. In particular, if a_t, L_t do not depend on t, then A generates a strongly continuous semigroup on $C_{\psi_2,\infty}$ with C_{ψ_1} being an invariant core.

Proof. The differentiability of $S^{t,r}\phi(x)$ for each x follows from (C.8) (better to say its time reversal version). Differentiating equation (C.10) one sees directly that $S^{t,r}\phi$ satisfies (C.9) and al required formulas hold point-wise. To show that they hold in the topology of $B_{\psi_2,\infty}$ one needs to show that the operators $A_t(\phi)$ are continuous as functions from t to $B_{\psi_2,\infty}$ for each $\phi \in B_{\psi_1}$. But this follows directly from our continuity assumptions on a_t and L_t.

To show that the space C_{ψ_1} is invariant (and this wold obviously imply all other remaining statements), we shall approximate $S^{t,r}$ by the evolutions with bounded intensities. Let χ_x be a measurable function $X \mapsto [0,1]$ such that $\chi_n(x) = 1$ for $\psi_2(x) \leq n$ and $\chi_n(x) = 0$ for $\psi_2(x) \geq n+1$. Denote $\nu_t^n(x,dz) = \chi_n(x)\nu_t(x,dz)$, $a_t^n = \chi_n a_t$, and let $S_n^{t,r}$ (respectively A_t^n) denote the propagators constructed as in Proposition C.2 (respectively the operators from (C.1)) but with ν_t^n and a_t^n instead of ν_t and a_t. Then the propagators $S_n^{t,r}$ converge strongly in the Banach space $B_{\psi_2,\infty}$ to the propagator $S^{t,r}$. One can deduce this fact from a general statement on the convergence of propagators (see e.g. [26]), but a direct proof is even simpler. Namely, as $S^{t,r}$ and $S_n^{t,r}$ are uniformly bounded, it is enough to show the convergence for the elements ϕ of the invariant core B_{ψ_1}. For such a ϕ one has

$$(S^{t,r} - S_n^{t,r})(\phi) = \int_t^r \frac{d}{ds}S^{t,s}S_n^{s,r}\phi\, ds = \int_t^r S^{t,s}(A_s - A_s^n)S_n^{s,r}\phi\, ds, \tag{C.12}$$

where (C.11) was used. As by invariance $S_n^{s,r}\phi \in B_{\psi_1}$, it follows that $(A_s - A_s^n)S_n^{s,r}\phi \in B_{\psi_2}$ and tends to zero in the form of B_{ψ_2}, as $n \to \infty$, and hence the r.h.s. of (C.12) tends to zero in B_{ψ_2}, as $n \to \infty$.

To complete the proof it remains to observe that as the generators of $S_n^{t,r}$ are bounded, the corresponding semigroups preserves continuity (as they can be constructed as the convergent exponential series). Hence $S^{t,r}$ preserves the continuity as well, as $S^{t,r}\phi$ is a (uniform) limit of continuous functions.

Remark. Choosing $a_t = \|\nu_t(x,\cdot)\|$ and $\psi_1 = 1$ above yield a pure analytic construction of a strongly continuous propagator for a non-homogeneous jump type process. A more familiar probabilistic approach can be found e.g. in [5] (at least for the homogeneous case).

For our purposes we need a perturbed equation (C.9), namely the equation

$$\dot{u}_t = -(A_t - B_t)u_t, \quad u_r = \phi, \quad 0 \le t \le r, \tag{C.13}$$

where B_t are bounded operators in C_{ψ_1}, and its dual equation on measures, whose weak form is

$$\frac{d}{dt}(g, \xi_t) = ((A_t - B_t)g, \xi_t) \quad \xi_0 = \xi, \quad 0 \le t \le r, \tag{C.14}$$

i.e. has to hold for some class of test functions g. Motivated by the standard observation that formally equation (C.13) is equivalent to the integral equation

$$u_t = S^{t,r}\phi - \int_t^\tau S^{t,s} B_s u_s ds, \tag{C.15}$$

whose solution $u_t = U^{t,r}\phi$ one expects to obtain through the perturbation series

$$U^{t,r}\phi = S^{t,r}\phi - \int_t^r S^{t,s}B_s S^{s,r} ds\, \phi + \int_{t \le s_1 \le s_2 \le r} S^{t,s_1} B_{s_1} S^{s_1,s_2} B_{s_2} S^{s_2,r} ds_1 ds_2\, \phi + \cdots, \tag{C.16}$$

one arrives at the following result.

Theorem C.2 *Under the assumptions of Theorem C.1 suppose that $\psi_2(x) \to \infty$ as $x \to \infty$ and that a strongly continuous family of bounded operators $B_t : C_{\psi_2} \mapsto C_{\psi_1}$ is given. Then*

(i) series (C.16) *is absolutely convergent in $C_{\psi_2}(X)$ for any $\phi \in C_{\psi_2}(X)$ so that*

$$\|U^{t,r}\|_{C_{\psi_2}(X)} \le \|S^{t,r}\|_{C_{\psi_2}(X)} \exp\{(r-t) \sup_{t \le s \le r} \|B_s\|_{C_{\psi_2}(X)}\},$$

and defines a strongly continuous backward propagator $U^{t,r}$ in $C_{\psi_2,\infty}(X)$ with C_{ψ_1} being its invariant core (so that the analogues of (C.11) *hold);*

(ii) the operator $V^{r,s} = (U^{s,r})^*$ form a weakly continuous propagator in $\mathcal{M}_{\psi_2}$ yielding a unique (weakly continuous) solution to the Cauchy problem (C.14) *in the sense that it holds for all $g \in C_{\psi_1}(X)$;*

(iii) if f is an arbitrary continuous function tending to zero as $x \to \infty$, then the operators $V^{r,s} = (U^{r,s})^*$ are strongly continuous in the norm of $\mathcal{M}_{\psi_2 f}$ and solves a strong version of (C.14) *with derivative taken in the norm topology of $\mathcal{M}_{\psi_1 f}$.*

(iv) at last, if a family A_t^ω, B_t^ω of operators are given satisfying all the above conditions for each ω from an interval and such that $A_t^\omega - B_t^\omega$ depend strongly continuous on ω as operators $C_{\psi_1} \mapsto C_{\psi_2,\infty}$, then the corresponding resolving operators $U^{s,r}$ in $C_{\psi_2,\infty}$ depend strongly continuous on ω and their adjoint operators $V^{r,s}$ depend weakly continuous on ω in $\mathcal{M}_{\psi_2}$.

Proof. (i) (C.16) converges, because B_t are bounded. Other statements then follow directly from the corresponding facts about $S^{t,r}$.

(ii) The operators $V^{r,s}$ are weakly continuous in $\mathcal{M}_{\psi_2}(X)$ just because they are adjoint to strongly continuous operators in $C_{\psi_2,\infty}$. Next, the analogue of the third equation in (C.11) for $U^{t,r}$ is the equation

$$\frac{d}{dr}U^{s,r}g = U^{s,r}(A_r - B_r)g$$

that holds in $C_{\psi_2,\infty}(X)$ for any $g \in C_{\psi_1}$ according to (i). Passing to the adjoint operators it implies

$$\frac{d}{dr}(g, V^{r,s}Y) = ((A_r - B_r)g, V^{r,s}Y)$$

showing that $V^{r,s}$ yield a solution to (C.14). To show the uniqueness we shall use the method for the reduction of the uniqueness problem to the existence of certain solutions of the adjoint problem, see e.g. [27] in the Hilbert space setting and time independent generators. Let $0 < a < b < r$, $\chi_{[a,b]}(s)$ be an indicator function of $[a,b]$, and $v \in C_{\psi_1}(X)$. As $U^{t,r}$ solve (C.13), the function

$$\phi_t = \int_t^r U^{s,r}\chi_{[a,b]}(s)v\,ds$$

solves the problem

$$\frac{d}{dt}\phi_t = -(A_t - B_t)\phi_t + \chi_{[a,b]}(t)v, \quad \phi_r = 0, \tag{C.17}$$

in the sense that ϕ_t is continuous and satisfies (C.17) everywhere with possible exception of two points, where its derivative is not continuous. Now, to prove uniqueness for(C.14) it is enough to show that its any solution with $\xi_0 = 0$ vanishes. Assume that ξ_t is a weakly continuous function in $\mathcal{M}_{\psi_2}(X)$ such that $\xi_0 = 0$ and (C.14) holds for all $g \in C_{\psi_1}$. Integration by parts, (C.14) and weak continuity of ξ_t imply that

$$0 = (\phi_t, \xi_t)\mid_{t=0}^r = \int_0^r [(\dot{\phi}_t, \xi_t) + ((A_t - B_t)\phi_t, \xi_t)]dt$$

whenever ϕ_t has a uniformly bounded derivatives in $C_{\psi_2,\infty}(X)$ apart from a finite number of points. Using (C.17) yields the equation

$$\int_b^a (v, \xi_t)\,dt = 0.$$

As it holds for arbitrary $0 < a < b < r$, $v \in C_{\psi_1}(X)$, it implies that $\xi_t = 0$.

(iii) From (C.14) it follows that

$$(g, \xi_r) - (g, \xi_s) = \int_s^r ((A_t - B_t)g, \xi_t)\,dt, \quad 0 \le s \le r, \tag{C.18}$$

which implies that ξ_t is an absolutely continuous function of t in the norm $\mathcal{M}_{\psi_1}(X)$. From boundedness of ξ_t in $\mathcal{M}_{\psi_2 f}(X)$ (that follows from weak continuity) and the weak continuity in $\mathcal{M}_{\psi_1}(X)$ it follows the continuity in $\mathcal{M}_{\psi_2 f}(X)$. At last, again from (C.18) one concludes that ξ_t is continuously differentiable in $\mathcal{M}_{\psi_1 f}(X)$.

(iv) This is straightforward. Namely, one compares $U^{r,s}$ for various ω by a formula similar to (C.12). This yields the continuous dependence of $U^{r,s}\phi$ on ω for $\phi \in C^{\psi_1}(X)$. By approximation one extends this result to all $\phi \in C_\infty^{\psi_2}$.

Acknowledgments. The author is grateful to S. Gaubert and M. Akian for their hospitality in INRIA (France), the major part of this work being done during the author's visit to INRIA in the spring 2006 and to Martine Verneuille for the excellent typing of this manuscript in Latex.

References

[1] 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.

[2] V.P. Belavkin. A Quantum Nonadapted Ito Formula and Stochastic Analysis in Fock Scale. J. Funct. Anal. **102:2** (1991), 414-447.

[3] 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.

[4] P. Brémaud. Point Processes and Queues, Springer, 1981.

[5] Mu Fa Chen. From Markov Chains to Non-Equilibrium Particle Systems. World Scientific, 1992.

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

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

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

[9] A.M. Etheridge. An Introduction to Superprocesses. University Lecture Series, v. **20**. AMS Providence, 2000.

[10] A. Joffe, M. Métivier. Weak convergence of sequence of semimatingales with applications to multitype branching processes. Adv. Appl. Probab. **18** (1986), 20-65.

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

[12] M. Freidlin. Functional Integration and Partial Differential Equations. Princeton Univ. Press, Princeton, NY 1985.

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

[14] O. Kallenberg. Foundations of Modern Probability. Second ed., Springer 2002.

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

[16] 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.

[17] 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.

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

[19] 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), 9-38.

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

[21] V. N. Kolokoltsov. Nonlinear Markov Semigroups and Interacting Lévy Type Processes. Journ. Stat. Physics **126:3** (2007), 585-642.

[22] 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.

[23] 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.

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

[25] R.H. Martin. Nonlinear operators and differential equations in Banach spaces. New York, 1976.

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

[27] V.P. Maslov. Méthodes Opératorielles. Moscow, Nauka 1974 (in Russian). French transl. Moscow, Mir, 1987.

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

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

[30] M. Métivier. Weak convergence of mesaure-valued processes using Sobolev imbedding techniques. Proceedings 'Stochastic Partial Differential Equations', Trento 1985, Springer LNM 1236, 172-183.

[31] I. Mitoma. Tightness of probabilities on $C([0,1];\mathcal{S}')$ and $D([0,1];\mathcal{S}')$. Ann. Probab. **11:4** (1983), 989–999.

[32] I. Mitoma. An ∞-dimensional inhomogeneous Langevin's equation. J. Functional Analysis **61** (1985), 342-359.

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

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

[35] A. Ja. Povzner. The Boltzmann equation in the kinetic theory of gases. Mat. Sbornik **58** (1962), 65-86.

[36] R. Rebolledo. Sur l'existence de solution á certain problèmes de semimartingales, C.R. Acad. Sci. Paris, Ser. A-B **290:18** (1980), 843-846.