

Existence of Solutions to Path-Dependent Kinetic Equations and Related Forward-Backward Systems

Vassili Kolokoltsov^{1*}, Wei Yang²

¹Department of Statistics, University of Warwick, Coventry, UK

²Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK

Email: v.kolokoltsov@warwick.ac.uk, w.yang@strath.ac.uk

Received March 29, 2013; revised May 3, 2013; accepted May 16, 2013

Copyright © 2013 Vassili Kolokoltsov, Wei Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

This paper is devoted to path-dependent kinetics equations arising, in particular, from the analysis of the coupled backward-forward systems of equations of mean field games. We present local well-posedness, global existence and some regularity results for these equations.

Keywords: Kinetic Equation; Mean Field Control; Global Existence; Path Dependence; Nonlinear Markov Process; Coupled Backward-Forward Systems

1. Introduction

For a Banach space \mathbf{B} we denote by \mathbf{B}^* the dual Banach space of \mathbf{B} . The pairing between $f \in \mathbf{B}$ and $\mu \in \mathbf{B}^*$ is denoted by (f, μ) . The norm in \mathbf{B}^* is defined by $\|\mu\|_{\mathbf{B}^*} = \sup_{\|f\|_{\mathbf{B}} \leq 1} |(f, \mu)|$. For a $T > 0$, we denote by $C([0, T], \mathbf{B})$ the Banach space of continuous curves $\eta : [0, T] \rightarrow \mathbf{B}$ equipped with the norm $\sup_{t \in [0, T]} \|\eta_t\|_{\mathbf{B}}$.

A deterministic dynamic in \mathbf{B}^* can be naturally specified by a vector-valued ordinary differential equation

$$\dot{\mu}_t = \Psi(t, \mu_t) \quad (1)$$

with a given initial value $\mu \in \mathbf{B}^*$, where the mapping $(t, \eta) \mapsto \Psi(t, \eta)$ is from $\mathbf{R}^+ \times \mathbf{B}^*$ to \mathbf{B}^* . More generally, one often meets the situations when $\dot{\mu}$ does not belong to \mathbf{B}^* , but to some its extension. Namely, let \mathbf{D} be a dense subset of \mathbf{B} , which is itself a Banach space with the norm $\|\cdot\|_{\mathbf{D}} \geq \|\cdot\|_{\mathbf{B}}$. A deterministic dynamic in \mathbf{B}^* can be specified by Equation (1), where the mapping $(t, \eta) \mapsto \Psi(t, \eta)$ is from $\mathbf{R}^+ \times \mathbf{B}^*$ to \mathbf{D}^* . Written in weak form, Equation (1) means that, for all $f \in \mathbf{D}$,

$$(f, \dot{\mu}_t) = (f, \Psi(t, \mu_t)). \quad (2)$$

In many applications, Equation (2) appears in the form

*Supported by IPI RAN grants RFBR 11-01-12026 and 12-07-00115, and by the grant 4402 of the Ministry of Education and Science of Russia.

$$\frac{d}{dt}(f, \mu_t) = (A[t, \mu_t]f, \mu_t), \mu_0 = \mu, \quad (3)$$

where the mapping $(t, \eta) \mapsto A[t, \eta]$ is from $\mathbf{R}^+ \times \mathbf{B}^*$ to bounded linear operators $A[t, \eta]: \mathbf{D} \mapsto \mathbf{B}$ such that, for each pair $(t, \eta) \in \mathbf{R}^+ \times \mathbf{B}^*$, $A[t, \eta]$ generates a strongly continuous semigroup in \mathbf{B} . Of major interest is the case when \mathbf{B}^* is the space of measures on a locally compact space. It turns out that, in this case and under mild technical assumptions, an evolution (2) preserving positivity has to be of form (3) with the operators $A[t, \eta]$ generating Feller processes, see Theorems 6.8.1 and 11.5.1 from [1].

Equation (3) contains most of the basic equations from non-equilibrium statistical mechanics and evolutionary biology, see monograph [1] for an extensive discussion.

In this paper we are mostly interested in yet more general equation. Namely, let \mathcal{M} be a closed convex subset of \mathbf{B}^* , which is also closed in \mathbf{D}^* . For a $T > 0$, let $C([0, T], \mathcal{M}(\mathbf{D}^*))$ denote a closed convex subset of $C([0, T], \mathbf{D}^*)$ consisting of curves with values in \mathcal{M} , and $C_{\mu}([0, T], \mathcal{M}(\mathbf{D}^*))$ a closed convex subset of $C([0, T], \mathcal{M}(\mathbf{D}^*))$, consisting of curves $\{\eta_t\}$ with initial data $\eta_0 = \mu \in \mathcal{M}$.

The main object of this paper is a “path-dependent” version of Equation (3), that is

$$\frac{d}{dt}(f, \mu_t) = (A[t, \{\mu_s\}_{0 \leq s \leq t}]f, \mu_t), \mu_0 = \mu, \quad (4)$$

where

$$(t, \{\eta_s\}_{0 \leq s \leq T}) \mapsto A[t, \{\eta_s\}_{0 \leq s \leq T}]$$

maps $\mathbf{R}^+ \times C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$ to bounded linear operators $\mathbf{D} \mapsto \mathbf{B}$. We refer to equation (4) as the general *path-dependent kinetic equation*. It should hold for all test functions $f \in \mathbf{D}$. Compared to Equation (4), Equation (3) is often referred to as a *path-independent* case.

When the operators A only depend on the *history* of the trajectory of $\{\mu\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$, that is

$$\frac{d}{dt}(f, \mu_t) = (A[t, \{\mu_{\leq t}\}]f, \mu_t), \mu_0 = \mu, \quad (5)$$

we call (5) an *adapted kinetic equation*, where $\{\mu_{\leq t}\}$ is a shorthand for $\{\mu_s\}_{0 \leq s \leq t}$. Adapted kinetic equations can be seen as analytic analogs of stochastic differential equations with adapted coefficients, and their well-posedness can be obtained by similar methods. When the generators A only depend on the *future* of the trajectory of $\{\mu\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$, that is

$$\frac{d}{dt}(f, \mu_t) = (A[t, \{\mu_{\geq t}\}]f, \mu_t), \mu_0 = \mu, \quad (6)$$

we call (6) an *anticipating kinetic equation*, where $\{\mu_{\geq t}\}$ is a shorthand for $\{\mu_s\}_{t \leq s \leq T}$.

Equation (1.4) has many applications. Let us briefly explain the crucial role played by this equation in the mean field game (MFG) methodology, which is based on the analysis of coupled systems of forward-backward evolutions and which constitutes a quickly developing area of research in modern theory of optimization, see detail e.g. in [2-4].

Assume that the objective of an agent described by a controlled stochastic process $X(s)$ (passing through x at time t), given an evolution $\hat{\mu}$ of the empirical distributions of a large number of other players, is to maximize (over a suitable class of controls $\{u\}$) the payoff

$$\begin{aligned} & V(t, x, \hat{\mu}_{\geq t}, u_{\geq t}) \\ &= \mathbb{E} \left[\int_t^T J(s, X(s), \hat{\mu}_s, u_s) ds + V^T(X(T)) \right], \end{aligned}$$

By dynamic programming the optimal payoff $V(t, x, \hat{\mu}_{\geq t})$ of such an agent, which equals $\sup_u V(t, x, \hat{\mu}_{\geq t}, u_{\geq t})$ should satisfy certain HJB equation (backward evolution). On the other hand, when all optimal controls $\{u_t = u_t(\hat{\mu}_{\geq t})\}$ are found, the empirical measure μ of the resulting process satisfies the controlled kinetic equation of type (3) (forward equation), that is

$$\frac{d}{dt}(f, \mu_t) = (A[t, \mu_t, u_t(\hat{\mu}_{\geq t})]f, \mu_t), \mu_0 = \mu. \quad (7)$$

The main consistency condition of MFG is in the requirement that the initial $\hat{\mu}$ coincides with the resulting

μ . Equalizing $\hat{\mu} = \mu$ in (7) clearly leads to anticipating kinetic equation of type (6).

Our main results concern the well-posedness of adaptive kinetic Equations (5), the local well-posedness and global existence of anticipating and general path dependent kinetic equations and finally some regularity result for path-independent equations arising from their probabilistic interpretations. This yield an improved version of the existence results of the unpublished preprint [4].

2. Main Results

Let us recall the notion of propagators needed for the proper formulation of our results.

For a set S , a family of mappings $U^{t,r}$ from S to itself, parametrized by the pairs of numbers $r \leq t$ (resp. $r \geq t$) from a given finite or infinite interval is called a (*forward*) *propagator* (resp. a *backward propagator*) in S , if $U^{t,t}$ is the identity operator in S for all t and the following *chain rule*, or *propagator equation*, holds for $r \leq s \leq t$ (resp. for $t \leq s \leq r$):

$$U^{t,s}U^{s,r} = U^{t,r}.$$

A backward propagator $U^{t,r}$ of bounded linear operators on a Banach space \mathbf{B} is called strongly continuous if the operators $U^{t,r}$ depend strongly continuously on t and r .

Suppose $U^{t,r}$ is a strongly continuous backward propagator of bounded linear operators on a Banach space with a common invariant domain \mathbf{D} . Let $A_t, t \geq 0$, be a family of bounded linear operators $\mathbf{D} \mapsto \mathbf{B}$ that are strongly continuous in t outside a set S of zero-measure in \mathbf{R} . Let us say that the family A_t *generates* $U^{t,r}$ on \mathbf{D} if, for any $f \in \mathbf{D}$, the equations

$$\begin{aligned} \frac{d}{ds} U^{t,s} f &= U^{t,s} A_s f, \\ \frac{d}{ds} U^{s,r} f &= -A_s U^{s,r} f, \quad 0 \leq t \leq s \leq r, \end{aligned} \quad (8)$$

hold for all s outside S with the derivatives taken in the topology of \mathbf{B} . In particular, if the operators A_t depend strongly continuously on t , equations (8) hold for all s and $f \in \mathbf{D}$, where for $s = t$ (resp. $s = r$) it is assumed to be only a right (resp. left) derivative. In the case of propagators in the space of measures, the second equation in (8) is called *the backward Kolomogorov equation*.

We can now formulate our main results.

Theorem 2.1 (local well-posedness for general ‘‘path-dependent’’ case) *Let \mathcal{M} be a bounded convex subset of \mathbf{B}^* with $\sup_{\mu \in \mathcal{M}} \|\mu\|_{\mathbf{B}^*} \leq K$, which is closed in the norm topologies of both \mathbf{B}^* and \mathbf{D}^* . Suppose that*

1) the linear operators $A[t, \{\xi\}]: \mathbf{D} \mapsto \mathbf{B}$ are uniformly bounded and Lipschitz in $\{\xi\}$, i.e. for any

$$\begin{aligned} & \{\xi\}, \{\eta\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*)) \\ & \sup_{t \in [0, T]} \|A[t, \{\xi\}] - A[t, \{\eta\}]\|_{\mathbf{D} \rightarrow \mathbf{B}} \\ & \leq c_1 \sup_{t \in [0, T]} \|\xi_t - \eta_t\|_{\mathbf{D}^*}, \end{aligned} \quad (9)$$

$$\sup_{t \in [0, T]} \|A[t, \{\xi\}]\|_{\mathbf{D} \rightarrow \mathbf{B}} \leq c_1 \quad (10)$$

for a positive constant c_1 ;

2) for any $\{\xi\} \in C_\mu([0, T], (\mathbf{D}^*))$, let the operator curve $A[t, \{\xi\}]: \mathbf{D} \mapsto \mathbf{B}$ generate a strongly continuous backward propagator of bounded linear operators $U^{t,s}[\{\xi\}]$ in \mathbf{B} , $0 \leq t \leq s$, on the common invariant domain \mathbf{D} , such that

$$\begin{aligned} & \|U^{t,s}[\{\xi\}]\|_{\mathbf{D} \rightarrow \mathbf{D}} \leq c_2 \text{ and} \\ & \|U^{t,s}[\{\xi\}]\|_{\mathbf{B} \rightarrow \mathbf{B}} \leq c_3, t \leq s, \end{aligned} \quad (11)$$

for some positive constants c_2, c_3 , and with their dual propagators $\tilde{U}^{s,t}[\{\xi\}]$ preserving the set \mathcal{M} .

Then, if

$$c_1 c_2 c_3 K T < 1, \quad (12)$$

the Cauchy problem (4) is well posed, that is for any $\mu \in \mathcal{M}$, it has a unique solution $\Phi^t(\mu) \in \mathcal{M}$ (that is (4) holds for all $f \in \mathbf{D}$) that depends Lipschitz continuously on time t and the initial data in the norm of \mathbf{D}^* , i.e.

$$\begin{aligned} & \|\Phi^s(\mu) - \Phi^t(\mu)\|_{\mathbf{D}^*} \leq c_1 c_2 (s - t), \\ & 0 \leq t \leq s \leq T, \end{aligned} \quad (13)$$

and for $\mu, \eta \in \mathcal{M}$

$$\begin{aligned} & \|\{\Phi^t(\mu)\} - \{\Phi^t(\eta)\}\|_{C([0, T], \mathcal{M}(\mathbf{D}^*))} \\ & \leq \frac{c_2}{1 - c_1 c_2 c_3 K T} \|\mu - \eta\|_{\mathbf{D}^*}. \end{aligned} \quad (14)$$

Theorem 2.2 (global wellposedness for an “adapted” case) *Under the assumptions in Theorem 2.1, but without the locality constraint (12), the Cauchy problem (5) is well posed in \mathcal{M} and its unique solution depends Lipschitz continuously on initial data in the norm of \mathbf{D}^* .*

Theorem 2.3 (global existence of the solution for general “path dependent” case) *Under the assumptions in Theorem 2.1, but without the locality constraint (12), assume additionally that for any t from a dense subset of $[0, T]$, the set*

$$\left\{ \tilde{U}^{t,0}[\{\xi\}]\mu : \{\xi\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*)) \right\} \quad (15)$$

is relatively compact in \mathcal{M} . Then a solution to the Cauchy problem (4) exists in \mathcal{M} .

In Proposition 4.3 in Section 4, we give the conditions under which the compactness assumption (15) holds.

3. Proofs of the Main Results

Proof of Theorem 2.1

$$\begin{aligned} & \text{By duality, for any } \{\xi^1\}, \{\xi^2\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*)) \\ & \left(f, \left(\tilde{U}^{t,0}[\{\xi^1\}] - \tilde{U}^{t,0}[\{\xi^2\}] \right) \mu \right) \\ & = \left(\left(U^{0,t}[\{\xi^1\}] - U^{0,t}[\{\xi^2\}] \right) f, \mu \right). \end{aligned}$$

By (8),

$$\begin{aligned} & U^{0,t}[\{\xi^2\}] - U^{0,t}[\{\xi^1\}] = U^{0,r}[\{\xi^2\}] U^{r,t}[\{\xi^1\}] \Big|_{r=0}^t \\ & = \int_0^t \frac{d}{dr} \left(U^{0,r}[\{\xi^2\}] U^{r,t}[\{\xi^1\}] \right) dr \\ & = \int_0^t U^{0,r}[\{\xi^2\}] \left(A[r, \{\xi^2\}] - A[r, \{\xi^1\}] \right) U^{r,t}[\{\xi^1\}] ds. \end{aligned}$$

Then, together with assumptions (9) and (11),

$$\begin{aligned} & \left\| \left(\tilde{U}^{t,0}[\{\xi^1\}] - \tilde{U}^{t,0}[\{\xi^2\}] \right) \mu \right\|_{\mathbf{D}^*} \\ & \leq \left\| U^{0,t}[\{\xi^1\}] - U^{0,t}[\{\xi^2\}] \right\|_{\mathbf{D} \rightarrow \mathbf{B}} \|\mu\|_{\mathbf{B}^*} \\ & \leq c_1 c_2 c_3 t K \left\| \{\xi^1\} - \{\xi^2\} \right\|_{C([0, T], \mathcal{M}(\mathbf{D}^*))}. \end{aligned} \quad (16)$$

Consequently, if (12) holds, the mapping $\{\xi\} \mapsto \{\tilde{U}^{t,0}[\{\xi\}]\}_{t \in [0, T]}$ is a contraction in

$C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$. Hence by the contraction principle there exists a unique fixed point for this mapping and hence a unique solution to Equation (4).

Inequality (13) follows directly from (4). Finally, if $\Phi^t(\mu) = \mu_t$ and $\Phi^t(\eta) = \eta_t$, then

$$\begin{aligned} & \mu_t - \eta_t = \tilde{U}^{t,0}[\{\mu\}]\mu - \tilde{U}^{t,0}[\{\eta\}]\eta \\ & = \left(\tilde{U}^{t,0}[\{\mu\}] - \tilde{U}^{t,0}[\{\eta\}] \right) \mu + \tilde{U}^{t,0}[\{\eta\}](\mu - \eta). \end{aligned}$$

From (11) and (16),

$$\begin{aligned} & \|\{\mu_t\} - \{\eta_t\}\|_{C([0, T], \mathcal{M}(\mathbf{D}^*))} \\ & \leq c_1 c_2 c_3 T K \|\{\mu\} - \{\eta\}\|_{C([0, T], \mathcal{M}(\mathbf{D}^*))} + c_2 \|\mu - \eta\|_{\mathbf{D}^*} \end{aligned} \quad (17)$$

implying (14).

Proof of Theorem 2.2

For a $\mu \in \mathcal{M}$, let us construct an approximating sequence $\{\xi^n\} \in C([0, T], \mathcal{M}(\mathbf{D}^*))$, $n = 0, 1, \dots$, by defining $\xi_r^0 = \mu$ for $t \in [0, T]$ and then recursively $\xi_r^n = \tilde{U}^{t,0}[\xi_{s,t}^{n-1}]\mu$. By non-anticipation, arguing as in the proof of (16) above, we first get the estimate

$$\begin{aligned} & \sup_{0 \leq r \leq t} \|\xi_r^1 - \xi_r^0\|_{\mathbf{D}^*} \leq c_1 c_2 c_3 K t \text{ and then recursively} \\ & \sup_{0 \leq r \leq t} \|\xi_r^n - \xi_r^{n-1}\|_{\mathbf{D}^*} \leq c_1 c_2 c_3 K \int_0^t \sup_{0 \leq r \leq s} \|\xi_r^{n-1} - \xi_r^{n-2}\|_{\mathbf{D}^*} ds \end{aligned}$$

that implies (by straightforward induction) that, for all $t \in [0, T]$,

$$\left\| \{\xi^n\} - \{\xi^{n-1}\} \right\|_{C([0,t], \mathcal{M}(\mathbf{D}^*))} \leq \frac{1}{n!} (c_1 c_2 c_3 K t)^n.$$

Hence, the partial sums on the r.h.s. of the obvious equation

$$\xi_{\leq t}^n = (\xi_{\leq t}^n - \xi_{\leq t}^{n-1}) + \dots + (\xi_{\leq t}^1 - \xi_{\leq t}^0) + \xi_{\leq t}^0$$

converge, and thus the sequence ξ^n converges in $C([0, T], \mathbf{D}^*)$. The limit is clearly a solution to (5).

Finally, let us assume that μ_t and η_t are some solutions with the initial conditions μ and η respectively. Instead of (17), we now get

$$\begin{aligned} & \left\| \{\mu\} - \{\eta\} \right\|_{C([0,t], \mathcal{M}(\mathbf{D}^*))} \\ & \leq c_1 c_2 c_3 K \int_0^t \left\| \{\mu\} - \{\eta\} \right\|_{C([0,s], \mathcal{M}(\mathbf{D}^*))} ds + c_2 \|\mu - \eta\|_{\mathbf{D}^*}. \end{aligned}$$

By Gronwall's lemma, this implies that

$$\left\| \{\mu\} - \{\eta\} \right\|_{C([0,t], \mathcal{M}(\mathbf{D}^*))} \text{ does not exceed}$$

$c_2 \|\mu - \eta\|_{\mathbf{D}^*} \exp\{c_1 c_2 c_3 K t\}$ yielding uniqueness and Lipschitz continuity of solutions with respect to initial data.

Proof of Theorem 2.3

Since \mathcal{M} is convex, the space $C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$ is also convex. Since the dual operators $\tilde{U}^{t,0}[\{\xi\}]$ preserve the set \mathcal{M} , for any $\{\xi\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$, the curve $\tilde{U}^{t,0}[\{\xi\}]\mu$ belongs to $C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$ as a function of t . Hence, the mapping

$$\{\xi\} \rightarrow \{\tilde{U}^{t,0}[\{\xi\}]\mu, t \in [0, T]\}$$

is from $C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$ to itself. Moreover, by (16), this mapping is Lipschitz continuous.

Denote

$$\hat{C} = \left\{ \left\{ \tilde{U}^{t,0}[\{\xi\}]\mu \right\} : \{\xi\} \in C_\mu([0, T], \mathcal{M}(\mathbf{D}^*)) \right\}.$$

Together with (13), the assumption that set (15) is compact in \mathcal{M} for any t from a dense subset of $[0, T]$ implies that the set \hat{C} is relatively compact in $\{C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))\}$ (by the Arzela-Ascoli Theorem).

Finally, by Schauder fixed point theorem, there exists a fixed point in $\hat{C} \subset C_\mu([0, T], \mathcal{M}(\mathbf{D}^*))$, which gives the existence of a solution to (4).

4. Nonlinear Markov Evolutions and Its Regularity

This section is designed to provide a probabilistic interpretation and, as a consequence, certain regularity properties for nonlinear Markov evolution μ_t solving ki-

netic Equation (3) in the case when $\mathbf{B} = C_\infty(\mathbf{R}^d)$ is the Banach space of bounded continuous functions f on \mathbf{R}^d with $\lim_{x \rightarrow \infty} f(x) = 0$, equipped with sup-norm and $\mathcal{M} = \mathcal{P}(\mathbf{R}^d)$ is the set of probability measures on \mathbf{R}^d , so that \mathbf{B}^* is the space of signed Borel measures on \mathbf{R}^d and $K = \sup_{\mu \in \mathcal{M}(\mathbf{R}^d)} \|\mu\|_{\mathbf{B}^*} = 1$. As a consequence,

we shall present a simple criterion for the main compactness assumption of Theorem 2.3.

We shall denote $C_{Lip}(\mathbf{R}^d)$ the Banach space of bounded Lipschitz continuous functions f on \mathbf{R}^d with the norm

$$\|f\|_{Lip} = \sup_x |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

and $C_\infty^1(\mathbf{R}^d)$ (resp. $C_\infty^2(\mathbf{R}^d)$) the Banach space of continuously differentiable functions f on \mathbf{R}^d such that f and the derivative f' belongs to $C_\infty(\mathbf{R}^d)$, equipped with the norm

$$\|f\|_{C_\infty^1(\mathbf{R}^d)} := \sup_x (|f(x)| + |f'(x)|)$$

resp. twice continuously differentiable with $f, f', f'' \in C_\infty(\mathbf{R}^d)$ and the norm

$$\|f\|_{C_\infty^2(\mathbf{R}^d)} := \sup_x (|f(x)| + |f'(x)| + |f''(x)|).$$

Let $\{A[t, \mu] : t \geq 0, \mu \in \mathcal{P}(\mathbf{R}^d)\}$ be a family of operators in $C_\infty(\mathbf{R}^d)$ of the Lévy-Khintchin type, that is

$$\begin{aligned} & A[t, \mu]f(z) \\ & = \frac{1}{2} (G(t, z, \mu) \nabla, \nabla) f(z) + (b(t, z, \mu), \nabla f(z)) \\ & \quad + \int (f(z+y) - f(z) - (\nabla f(z), y) \mathbf{1}_{B_1}(y)) \nu(t, z, \mu, dy), \end{aligned} \tag{18}$$

where ∇ denotes the gradient operator; for $(t, z, \mu) \in [0, T] \times \mathbf{R}^d \times \mathcal{P}(\mathbf{R}^d)$, $G(t, z, \mu)$ is a symmetric non-negative matrix, $b(t, z, \mu)$ is a vector, $\nu(t, z, \mu, \cdot)$ is a Lévy measure on \mathbf{R}^d , i.e.

$$\int_{\mathbf{R}^d} \min(1, |y|^2) \nu(t, z, \mu, dy) < \infty, \nu(t, z, \mu, \{0\}) = 0, \tag{19}$$

depending measurably on t, z, μ , and $\mathbf{1}_{B_1}$ denotes, as usual, the indicator function of the unit ball in \mathbf{R}^d . Assume that each operator (18) generates a Feller process with one and the same domain \mathbf{D} such that $C_\infty^2(\mathbf{R}^d) \subseteq \mathbf{D} \subseteq C_\infty^1(\mathbf{R}^d)$.

Proposition 4.1 *Suppose the assumptions of Theorem 2.2 are fulfilled with generators $A[t, \mu]$ of "path-independent" type (3) and a probability measure μ is given. Then there exists a family of processes $\{X_{s,t}^\mu : \mu \in \mathcal{P}(\mathbf{R}^d)\}$ defined on a certain filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathcal{P}(\mathbf{R}^d))$ such that $\mu_t = \mathcal{L}(X_{s,t}^\mu)$ solves the Cauchy problem for Equation*

(3) with initial condition μ and $\{X_{s,t}^\mu\}$ solves the nonlinear martingale problem, specified by the family $\{A[t, \mu]\}$, that is, for any $f \in \mathbf{D}$,

$$f(X_{s,t}^\mu) - \int_s^t A[\tau, \mathcal{L}(X_{s,\tau}^\mu)] f(X_{s,\tau}^\mu) d\tau, s \leq t \quad (20)$$

is a martingale.

By the assumptions of Theorem 2.2, a solution $\mu_t \in \mathcal{P}(\mathbf{R}^d)$ of Equation (3) with initial condition $\mu_s = \mu$ specifies a propagator $\tilde{U}^{t,r}[\mu]$, $s \leq r \leq t$, of linear transformations in \mathbf{B}^* , solving the Cauchy problems for equation

$$\frac{d}{dt}(f, \nu_t) = (A[t, \mu_t] f, \nu_t). \quad (21)$$

In its turn, for any $\nu \in \mathcal{P}(\mathbf{R}^d)$, Equation (21) specifies marginal distributions of a usual (linear) Markov process $\{X_{s,t}^\mu(\nu)\}$ in \mathbf{R}^d with the initial measure ν . Clearly, the process $\{X_{s,t}^\mu(\mu)\}$ is a solution to our martingale problem.

We shall refer to the family of processes constructed in Proposition 4.1 as to *nonlinear Markov process* generated by the family $A[t, \mu]$.

Using martingales allows us to prove the following useful regularity property for the solution of kinetic equations.

Proposition 4.2 *Suppose the assumptions of Theorem 2.2 are fulfilled for a kinetic equation of “path-independent” type (3) with generators $A[t, \mu]$ of type (18). Let $\{X_{s,t}^\mu\}$ denote a nonlinear Markov process constructed from the family of generators $A[t, \mu]$ by Proposition 4.1. Assume, for $p \in (0, 2]$ and $P > 0$, the following boundedness condition holds:*

$$\sup_{x \in \mathbf{R}^d, t \geq 0, \mu \in \mathcal{M}} \left\{ |G(t, x, \mu)|, |b(t, x, \mu)|, \int \min(|y|^2, |y|^p) \nu(t, x, \mu, dy) \right\} \leq P, \quad (22)$$

and $\int |x|^p \mu(dx) = p_\mu < \infty$ for the initial measure $\mu_s = \mu$.

Then the distributions $\mathcal{L}(X_{s,t}^\mu) = \Phi^{t,s}(\mu)$, solving the Cauchy problem for Equation (3) with initial condition μ_s have uniformly bounded p th moments, *i.e.*

$$\int |x|^p \Phi^{t,s}(\mu)(dx) \leq c(T, P)[1 + p_\mu], \quad (23)$$

and are $\frac{1}{2}$ -Hölder continuous with respect to t in the space $(C_{Lip}(\mathbf{R}^d))^*$, *i.e.*

$$\begin{aligned} & \left\| \Phi^{t_1,s}(\mu) - \Phi^{t_2,s}(\mu) \right\|_{(C_{Lip}(\mathbf{R}^d))^*} \\ & \leq c(T, P) \sqrt{|t_1 - t_2|}, \forall t_1, t_2 \geq s \geq 0, \end{aligned} \quad (24)$$

with a positive constant c .

Proof. For a fixed trajectory $\{\mu_t\}_{t \geq 0}$ with initial value μ , one can consider $\{X_{s,t}^\mu\}$ as a usual Markov process. Using the estimates for the moments of such processes from formula (5.61) of [5] (more precisely, its straightforward extension to time non-homogeneous case), one obtains from (22) that

$$\begin{aligned} & \mathbb{E} \left[\min \left(|X_{s,t}^\mu - \hat{x}|^2, |X_{s,t}^\mu - \hat{x}|^p \right) \middle| X_{s,s}^\mu = \hat{x} \right] \\ & \leq e^{C(T,P)(t-s)} - 1. \end{aligned} \quad (25)$$

This implies (23) and the estimate

$$\mathbb{E} \left(|X_{s,t}^\mu - \hat{x}| \middle| X_{s,s}^\mu = \hat{x} \right) \leq C(T, P) \sqrt{t-s}, \quad (26)$$

where constants $C(T, P)$ can have different values in various formulas above.

Since $\Phi^{t,s}(\mu)$ is the distribution law of the process $\{X_{s,t}^\mu\}$,

$$\begin{aligned} & \left\| \Phi^{s,t_1}(\mu) - \Phi^{s,t_2}(\mu) \right\|_{(C_{Lip}(\mathbf{R}^d))^*} \\ & \leq \sup_{\|\psi\|_{C_{Lip}(\mathbf{R}^d)} \leq 1} \mathbb{E} \left| \psi(X_{s,t_1}^\mu) - \psi(X_{s,t_2}^\mu) \right| \\ & \leq \mathbb{E} |X_{s,t_1}^\mu - X_{s,t_2}^\mu|. \end{aligned} \quad (27)$$

From (26), (27) and Markov property, we get (24) as required.

Remark 4.1 *For diffusions with $G=1$, (24) was proved in [3].*

Our main purpose for presenting Proposition 4.2 lies in the following corollary that follows from (23) and an observation that a set of probability laws on \mathbf{R}^d with a bounded p th moment, $p > 0$, is tight.

Proposition 4.3 *Under the assumptions of Theorem 2.1 for generators $A[t, \{\mu_t\}]$ of Lévy-Khintchin type (18), but without locality condition (12), suppose the boundedness condition (22) holds for some $p \in (0, 2]$ and $P > 0$. Then the compactness condition from Theorem 2.3 (stating that set (15) is compact in $\mathcal{P}(\mathbf{R}^d)$) holds for any initial measure μ with a finite moment of p th order.*

5. Basic Examples of Operators $A[t, \mu]$

We present here basic examples of generators that fit to our main Theorems and are relevant to the study of mean field games. The most nontrivial condition of Theorem 2.1 is 2).

The simplest examples are McKean-Vlasov diffusions defined by SDE

$$dX_t = b(t, X_t, \{\mu_t\}) dt + \sigma(t, X_t, \{\mu_t\}) dW_t,$$

with corresponding generator

$$\begin{aligned}
 & A[t, \{\mu\}] f(x) \\
 &= (b(t, x, \{\mu\}), \nabla f(x)) + \frac{1}{2} (\sigma^2(t, x, \{\mu\}), \nabla^2 f(x)), \\
 & f \in C_\infty^2(\mathbf{R}^d),
 \end{aligned}$$

where the condition of Theorem 2.1 2) is known to follow, for Lipschitz continuous coefficients, from Ito's calculus.

Another example is supplied by *nonlinear Lévy processes* that are specified by generators of type (18) such that all coefficients do not depend on z , *i.e.*

$$\begin{aligned}
 & A[t, \{\mu\}] f(x) \\
 &= \frac{1}{2} (G(t, \{\mu\}) \nabla, \nabla) f(x) + (b(t, \{\mu\}), \nabla f)(x) \\
 &+ \int [f(x+y) - f(x) - (y, \nabla f(x)) \mathbf{1}_{B_1}(y)] \nu(t, \{\mu\}, dy).
 \end{aligned}$$

It follows from Proposition 7.1 of [1] that if the coefficients G, b, ν are continuous in t and Lipschitz continuous in $\{\mu\}$ in the norm of Banach space

$$C([0, T], C_\infty^2(\mathbf{R}^d))^*, \text{ i.e.}$$

$$\begin{aligned}
 & \|G(t, \{\mu\}) - G(t, \{\eta\})\| \\
 &+ \int \min(1, |y|^2) |\nu(t, \{\mu\}, dy) - \nu(t, \{\eta\}, dy)| \\
 &+ \|b(t, \{\mu\}) - b(t, \{\eta\})\| \\
 &\leq c \sup_{t \in [0, T]} \|\mu - \eta\|_{(C_\infty^2(\mathbf{R}^d))^*}, \forall t \in [0, T],
 \end{aligned}$$

with a positive constant c , then condition 2) of Theorem 2.1 holds with $\mathbf{D} = C_\infty^2(\mathbf{R}^d)$. Notice also that the usual examples of a functional F on measures given by monomials $\int g(x_1, \dots, x_n) \mu(dx_1) \dots \mu(dx_n)$ are Lipschitz continuous (or even smooth) in space $(C_\infty^2(\mathbf{R}^d))^*$ whenever g is sufficient smooth.

Another example is supplied by processes (describing lots of models including spatially homogeneous and mollified Boltzmann equation and interacting α -stable laws with $\alpha < 1$) with generators of order at most one

$$\begin{aligned}
 & A[t, \{\mu\}] f(x) \\
 &= (b(t, x, \{\mu\}), \nabla f(x)) \\
 &+ \int_{\mathbf{R}^d} (f(x+y) - f(x)) \nu(t, x, \{\mu\}, dy),
 \end{aligned}$$

with the Lévy measures ν having finite first moment $\int |y| \nu(t, x, \{\mu\}, dy)$. As is established in Theorem 4.17

of [1], if b, ν are continuous in t , Lipschitz continuous in x and Lipschitz continuous in μ , that is, $\forall t \in [0, T], x \in \mathbf{R}^d$,

$$\begin{aligned}
 & \|b(t, x, \{\mu\}) - b(t, x, \{\eta\})\| \\
 &+ \int |y| \nu(t, x, \{\mu\}, dy) - \nu(t, x, \{\eta\}, dy) \\
 &\leq c \sup_{t \in [0, T]} \|\{\mu_t - \eta_t\}\|_{(C_\infty^1(\mathbf{R}^d))^*},
 \end{aligned}$$

then condition 2) of Theorem 2.1 is satisfied with $\mathbf{D} = C_\infty^1(\mathbf{R}^d)$.

As other examples let us mention pure jump processes with bounded rates, where the conditions of Theorem 2.1 are satisfied with $\mathbf{D} = C_\infty^1(\mathbf{R}^d)$, and *nonlinear stable-like processes* (see [1]).

Let us note finally that not all interesting evolution of type (3) satisfy our Lipschitz continuity assumption. A different type of continuity should be applied for coefficients depending on measures via quantiles, e.g. value at risk (VAR), which is analyzed in [6] inspired by preprint [7].

REFERENCES

- [1] V. N. Kolokoltsov, "Nonlinear Markov Processes and Kinetic Equations. Cambridge Tracks in Mathematics 182," Cambridge University Press, Cambridge, 2010.
- [2] O. Guéant, J.-M. Lasry and P.-L. Lions, "Mean Field Games and Applications. Paris-Princeton Lectures on Mathematical Finance 2010," Springer, Berlin, pp. 205-266.
- [3] M. Huang, R. P. Malhamé and P. E. Caines, "Large Population Stochastic Dynamic Games: Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle," *Communications in Information and Systems*, Vol. 6, No. 3, 2006, pp. 221-252.
- [4] V. N. Kolokoltsov, J. J. Li and W. Yang, "Mean Field Games and Nonlinear Markov Processes," 2011. arXiv:1112.3744v2
- [5] V. N. Kolokoltsov, "Markov Processes, Semigroups and Generators," De Gruyter, 2011.
- [6] V. N. Kolokoltsov, "Nonlinear Diffusions and Stable-Like Processes with Coefficients Depending on the Median or VaR," *Applied Mathematics and Optimization*, 2012. <http://arxiv.org/abs/1207.5925>
- [7] D. Crisan, Th. Kurtz and Y. Lee, "Conditional Distributions, Exchangeable Particle Systems, and Stochastic Partial Differential Equations," Preprint, 2012.