On mean field games with common noise and McKean-Vlasov SPDEs

Vassili N. Kolokoltsov[1] and Marianna Troeva[2]

[1]*Department of Statistics, University of Warwick, Coventry CV4 7AL UK,*
Email: v.kolokoltsov@warwick.ac.uk
[2]*Research Institute of Mathematics, North-Eastern Federal University, 58 Belinskogo str.,*
Yakutsk 677000 Russia, Email: troeva@mail.ru

Abstract

We formulate the MFG limit for N interacting agents with a common noise as a single quasi-linear deterministic infinite-dimensional partial differential second order backward equation. We prove that any its (regular enough) solution provides an $1/N$-Nash-equilibrium profile for the initial N-player game. We use the method of stochastic characteristics to provide the link with the basic models of MFG with a major player. We develop two auxiliary theories of independent interest: sensitivity and regularity analysis for McKean-Vlasov SPDEs and the $1/N$-convergence rate for the propagation of chaos property of interacting diffusions.

1 Introduction

We shall denote by $\mathcal{M}^{sign}(\mathbf{R}^n)$ the space of signed Borel measures on $\mathbf{R}^n$ of finite total variation, by $\mathcal{M}(\mathbf{R}^n)$ its cone of positive measures, by $\mathcal{M}^{sign}_\lambda(\mathbf{R}^n)$, $\mathcal{M}_\lambda(\mathbf{R}^n)$ their subsets of total variation norm not exceeding λ, by $\mathcal{P}(\mathbf{R}^n)$ the set of probability measures. We shall use the standard notation $(\phi, \mu) = \int \phi(x)\mu(dx)$ for the pairing of functions and measures. By $\mathbf{E}$ we denote the expectation.

Let us consider N agents, whose positions are governed by the system of SDEs

$$dX_t^i = b(t, X_t^i, \mu_t^N, u_t^i)\, dt + \sigma_{ind}(X_t^i)dB_t^i + \sigma_{com}(X_t^i)dW_t, \tag{1}$$

where all X_t^i belong to $\mathbf{R}$, $W_t, B_t^1, \cdots, B_t^N$ are independent one-dimensional standard Brownian motions, W_t, referred to as the common noise, and all B_t^j, referred to as the idiosyncratic or individual noises, the subscripts 'com' and 'ind' referred to the objects related to the common or to the individual noises. The parameters $u_t^i \in U \subset \mathbf{R}^m$ are controls available to the players, trying to minimize their payoffs

$$V_{[t,T]}^i(x) = \mathbf{E}\left[\int_t^T J(s, X_s^i, \mu_s^N, u_s^i)\, ds + V_T(X_T^i, \mu_T^N)\right], \tag{2}$$

depending on the action of other players, with the given functions J and V_T. The coefficient $b(t, x, \mu, u)$ is a function of $t \in \mathbf{R}, x \in \mathbf{R}, u \in U$ and a measure $\mu \in \mathcal{M}^{sign}(\mathbf{R})$, and μ_t^N in (1) is

$$\mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}.$$

In general, the function b needs to be defined only for μ from the set of probability measures $\mathcal{P}(R^n)$. However, to use smoothness with respect to μ it is convenient (though not necessary) to have this function defined on a larger space. In the usual examples, b depend on μ via a finite set of moments of type

$$F_j(\mu) = \int \tilde{F}_j(x_1, \cdots, x_k) \mu(dx_1) \cdots \mu(dx_k), \tag{3}$$

with some bounded measurable symmetric functions $\tilde{F}_j$.

For simplicity, we shall assume b to be linear in u, that is,

$$b(t, x, \mu, u) = b_1(t, x, \mu) + b_2(t, x, \mu)u, \tag{4}$$

though other weaker assumptions are possible.

Remark 1. *1) For this paper we shall stick to a smooth dependence of b on μ. However, more singular dependencies are also of interest, for instance, the dependence on μ via its quantile, see [1]. This case will be discussed in our subsequent publication based on the theory of SDEs with coefficients depending on quantiles developed in [2]. 2) In an attempt to present our main result in the most clear way, we make several simplifying assumptions, primarily that all objects are one dimensional and that σ_{ind} does not depend on μ and u, which can be relaxed causing the increase of technicalities. 3) We consider the simplest common noise W_t. It would be natural to extend the theory to the space-time white noise $W(dxdt)$ or even to a more general noise expressed in terms of functional semimartingales $F(x, t)$ analyzed in [3].*

It is known (see e.g. [4]) that, for fixed common functions $u_t^i(X_t^i) = u_t(X_t^i)$, and under appropriate regularity assumptions on $b, \sigma_{ind}, \sigma_{com}$ the system (1) is well-posed and the corresponding empirical measures μ_t^N converge, as $N \to \infty$, to the unique solution μ_t of the nonlinear SPDE of the McKean-Vlasov type

$$d(\phi, \mu_t) = (L[t, \mu_t, u_t]\phi, \mu_t) \, dt + (\sigma_{com}(.)\nabla\phi, \mu_t) \, dW_t, \tag{5}$$

which is written here in the weak form meaning that it should hold for all $\phi \in C^2(\mathbf{R})$, and where ∇ is the derivative with respect to the space variable x and

$$L[t, \mu_t, u_t]\phi(x) = \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\frac{\partial^2 \phi}{\partial x^2} + b(t, x, \mu_t, u_t(x))\frac{\partial \phi}{\partial x}. \tag{6}$$

Let us mention directly that in our approach it is more convenient to work with the Stratonovich differentials. Namely, by the usual rule $Y \circ dX = Y dX + \frac{1}{2} dY dX$, equation (5) rewrites in the Stratonovich form as

$$d(\phi, \mu_t) = (L_{St}[t, \mu_t, u_t]\phi, \mu_t) \, dt + (\sigma_{com}(.)\nabla\phi, \mu_t) \circ dW_t, \tag{7}$$

with

$$L_{St}[t, \mu_t, u_t]\phi(x) = \frac{1}{2}\sigma_{ind}^2(x)\frac{\partial^2 \phi}{\partial x^2} + [b(t, x, \mu_t, u_t(x)) - \frac{1}{2}\sigma_{com}\sigma'_{com}(x)]\frac{\partial \phi}{\partial x}. \qquad (8)$$

With some abuse of notation, we shall often identify measures with their densities (whenever they exist) with respect to Lebesgue measures, thus writing the strong form of equation (5) as

$$d\mu_t = L'[t, \mu_t, u_t]\mu_t \, dt - \nabla(\sigma_{com}(.)\mu_t) \, dW_t, \qquad (9)$$

with

$$L'[t, \mu, u_t]\mu = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left[(\sigma_{ind}^2 + \sigma_{com}^2)\mu)\right] - \frac{\partial}{\partial x}[b(t, x, \mu, u_t)\mu]. \qquad (10)$$

This identification does not cause ambiguity, because under non-degeneracy of $\sigma_{ind}^2 + \sigma_{com}^2$ that we shall always assume, any solution $\mu_t[\mu_0]$ to (5) has a density with respect to Lebesgue measure at any $t > 0$, even if μ_0 does not.

Recall now that the optimal control problem facing each player, say X_t^1, is to minimize cost (2). Now the crucial difference with the games without common noise starts to reveal itself. For games without noise, one expects to get a deterministic curve μ_t in the limit of large N, so that in the limit, each player should solve a usual optimization problem for a diffusion in **R**. Here the limit is stochastic, and thus even in the limit the optimization problem faced by each player is an optimization with respect to an infinite-dimensional, in fact measure-valued, process.

In fact, for fixed N, if all players, apart from the first one, are using the same control $u_{com}(t, x, \mu)$, the optimal payoff for the first player is found from the HJB equation for the diffusion governed by (1), that is, the HJB equation (where we denote X^1 by x),

$$\frac{\partial V}{\partial t}(t, x, \mu) + \inf_u \left[b(t, x, \mu_t, u)\frac{\partial V}{\partial x} + J(t, x, \mu_t, u)\right] + \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\frac{\partial^2 V}{\partial x^2}$$

$$+ \sum_{j \neq 1} b(t, x_j, \mu_t, u_{com}(t, x_j, \mu))\frac{\partial V}{\partial x_j} + \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x_j)\frac{\partial^2 V}{\partial x_j^2}$$

$$+ \sum_{j \neq 1} \sigma_{com}(x)\sigma_{com}(x_j)\frac{\partial^2 V}{\partial x_1 \partial x_j} + \sum_{1 < i < j} \sigma_{com}(x_i)\sigma_{com}(x_j)\frac{\partial^2 V}{\partial x_i \partial x_j} = 0. \qquad (11)$$

As will be shown, in the limit when $(\delta_{x_1} + \cdots + \delta_{x_N})/N$ converge to the process μ_t, this equation turns into the limiting HJB equation

$$\frac{\partial V}{\partial t}(t, x, \mu) + \inf_u \left[b(t, x, \mu, u)\frac{\partial V}{\partial x} + J(t, x, \mu, u)\right] + \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\frac{\partial^2 V}{\partial x^2}$$

$$+ \Lambda_{lim}V(t, x, \mu) + \int \sigma_{com}(x)\sigma_{com}(y)\frac{\partial^2}{\partial x \partial y}\frac{\delta V(t, x, \mu)}{\delta \mu(y)}\mu(dy) = 0, \qquad (12)$$

where the operator Λ_{lim} is calculated in (31) with u_{com} as the control.

If J is convex, the infimum here is achieved on the single point

$$\hat{u}_{ind}(t, x, \mu) = argmin_u \left[b(t, x, \mu_t, u)\frac{\partial V}{\partial x} + J(t, x, \mu_t, u)\right]$$

$$= -\left(\frac{\partial J}{\partial u}\right)^{-1}\left(b_2(t, x, \mu_t)\frac{\partial V}{\partial x}\right), \qquad (13)$$

3

where $\left(\dfrac{\partial J}{\partial u}\right)^{-1}$ is the inverse function to the function $\dfrac{\partial J}{\partial u}$.

Now the difference with the usual MFG is fully seen. Instead of a pair of coupled forward-backward equations we have now one single infinite-dimensional equation (12). Namely, for any curve $u_{com}(t, x, \mu)$ (defining Λ_{lim} in (31) and thus in (12)), we should solve equation (12) with a given terminal condition leading to the optimal control (13). The key *MFG consistency requirement* is now given by the equation

$$\hat{u}_{ind}(t, x, \mu) = u_{com}(t, x, \mu). \tag{14}$$

This can be interpreted as having a limiting game of two players, a tagged player and a measure-valued player, for which we are looking for a symmetric Nash equilbrium.

Equivalently, the MFG consistency (14) can be encoded into a single quasi-linear deterministic infinite-dimensional partial differential second order backward equation on the function $V(t, x, \mu)$, which we present now in full substituting Λ_{lim} from (31) and (14) in (12):

$$\frac{\partial V}{\partial t}(t, x, \mu) + \left[b(t, x, \mu, u)\frac{\partial V}{\partial x} + J(t, x, \mu, u)\right]\bigg|_{u(t,x,\mu)=-(\partial J/\partial u)^{-1}(b_2(t,x,\mu)\frac{\partial V}{\partial x})}$$

$$+\frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\int_{\mathbf{R}^2}\sigma_{com}(y)\sigma_{com}(z)\frac{\partial^2}{\partial y \partial z}\frac{\delta^2 V(t,x,\mu)}{\delta\mu(y)\delta\mu(z)}\mu(dy)\mu(dz)$$

$$+\int\left(\left[b(t, x, \mu, u(t,x,\mu))\right]\big|_{u(t,x,\mu)=-(\frac{\partial J}{\partial u})^{-1}(b_2(t,x,\mu)\frac{\partial V}{\partial z})}\nabla + \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\nabla^2\right]$$

$$\cdot\frac{\delta V(t,x,\mu)}{\delta\mu(.)}\right)(y)\mu(dy) + \int\sigma_{com}(x)\sigma_{com}(y)\frac{\partial^2}{\partial x \partial y}\frac{\delta V(t,x,\mu)}{\delta\mu(y)}\mu(dy) = 0, \tag{15}$$

with a given terminal condition

$$V(t, x, \mu)|_{t=T} = V_T(x, \mu), \quad \mu_t|_{t=0} = \mu_0. \tag{16}$$

If

$$J(t, x, \mu, u) = \frac{1}{2}u^2, \quad b_2(t, x, \mu) = 1,$$

then (15) simplifies to

$$\frac{\partial V}{\partial t}(t, x, \mu) + \left[b\left(t, x, \mu, -\frac{\partial V}{\partial x}\right)\frac{\partial V}{\partial x} + \frac{1}{2}\left(\frac{\partial V}{\partial x}\right)^2\right]$$

$$+\frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\frac{\partial^2 V}{\partial x^2} + \frac{1}{2}\int_{\mathbf{R}^2}\sigma_{com}(y)\sigma_{com}(z)\frac{\partial^2}{\partial y \partial z}\frac{\delta^2 V(t,x,\mu)}{\delta\mu(y)\delta\mu(z)}\mu(dy)\mu(dz)$$

$$+\int\left(\left[b(t, x, \mu, -\nabla V(t,x,\mu))\nabla + \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\nabla^2\right]\frac{\delta V(t,x,\mu)}{\delta\mu(.)}\right)(y)\mu(dy)$$

$$+\int\sigma_{com}(x)\sigma_{com}(y)\frac{\partial^2}{\partial x \partial y}\frac{\delta V(t,x,\mu)}{\delta\mu(y)}\mu(dy) = 0. \tag{17}$$

The MFG methodology suggests that for large N the optimal behavior of players arises from the control $\hat{u}$ given by (13) with V solving (15), or equivalently, satisfying the consistency condition (14) (see MFG3) below).

To justify this claim one is confronted essentially with the 3 problems:

MFG1): Prove well-posedness of (or at least the existence of the solution to) the problem (15) or (14);

MFG2): Analyze the Nash equilibria of the N-player game given by (1), (2) and prove that these equilibria (or at least their subsequence) converge, as $N \to \infty$, to a solution of the problem (15) or (14); assess the convergence rates;

This is a very difficult problem and only partial results subject to strong hypotheses have so far been obtained (see e.g. [5–7]).

MFG3): Show that a solution to the problem (15) or (14) provides a profile of symmetric strategies $\hat{u}_t(x)$, which is an ϵ-Nash equilibrium of the N-player game given by (1), (2) and the initial distribution of players μ_0, with $\epsilon(N) \to 0$ as $N \to \infty$; estimate the error-term $\epsilon(N)$.

Questions MFG2), MFG3) are of course two facets of the same problem on how well the solutions to the limiting problem ((5),(12),(13),(16)) approximate a finite player game, but the methods of dealing with these problems can be rather different.

To link with the usual MFG, let us notice that for the case without common noise given by (1) with $\sigma_{com} = 0$, equation (17), say, turns to

$$\frac{\partial V}{\partial t}(t, x, \mu) + \left[b\left(t, x, \mu, -\frac{\partial V}{\partial x}\right) \frac{\partial V}{\partial x} + \frac{1}{2}\left(\frac{\partial V}{\partial x}\right)^2 \right] + \frac{1}{2}\sigma_{ind}^2(x)\frac{\partial^2 V}{\partial x^2}$$

$$+ \int \left(\left[b(t, x, \mu, -\nabla V(t, x, \mu))\nabla + \frac{1}{2}\sigma_{ind}^2(.)\nabla^2 \right] \frac{\delta V(t, x, \mu)}{\delta \mu(.)} \right) (y)\mu(dy) = 0, \qquad (18)$$

giving a single-equation approach to usual MFG. In fact, solving this equation for a function $V(t, x, \mu)$ is equivalent to solving first the deterministic (forward) equation (9) with $\sigma_{com} = 0$ and then the backward equation

$$\frac{\partial V}{\partial t}(t, x, \mu_t) + \left[b\left(t, x, \mu_t, -\frac{\partial V}{\partial x}\right) \frac{\partial V}{\partial x} + \frac{1}{2}\left(\frac{\partial V}{\partial x}\right)^2 \right] + \frac{1}{2}\sigma_{ind}^2(x)\frac{\partial^2 V}{\partial x^2} = 0$$

for a function $V(t, x)$.

In a more abstract form the link between the forward-backward formulation and the single backward formulation is as follows. If (x_t, μ_t) is a controlled Markov process (not necessarily measure-valued), optimal payoff is defined via the corresponding HJB on a function $V(t, x, \mu)$ of three arguments (corresponds to our general common noise case). If the evolution of the coordinate μ_t is deterministic and does not depend on x and its control, one can (alternatively and equivalently) first solve this deterministic equation on μ (usual forward part of the basic MFG) and then substitute it in the basic HJB to get the equation on $V(t, x)$, the function of two arguments only, with μ_t included in the time dependence (usual backward part of the basic MFG). Thus obtained forward-backward system is then usually solved by the fixed point argument, see e.g. [8,9]. This decomposition into forward-backward system is not available in general.

In this paper we are going to concentrate exclusively on question MFG3), aiming at proving the error-estimate of order $\epsilon(N) \sim 1/N$. Our approach will be based on interpreting (by means of Ito's formula) the common noise as a kind of binary interaction of agents (in addition to the usual mean-field interaction of the standard situation without

common noise) and then reducing the problem to the sensitivity analysis for McKean-Vlasov SPDE.

The question MFG1) can be approached via the methods of papers [10, 11], which will be addressed in another publication. Some existence can be also derived from [12], which has however a slightly different formulation than the present one.

Our paper is organized as follows. Next section provides a short literature review. Section 3 provides our main results and their proofs. Sections 4-6 are devoted to the regularity and sensitivity analysis of the solutions to the McKean-Vlasov SPDEs and the related properties of the corresponding measure-valued Markov processes.

2   Brief literature review

Mean-field games present a quickly developing area of the game theory. It was initiated by Lasry-Lions [13] and Huang-Malhame-Caines [8, 14, 15], see [16–21] for recent surveys, as well as [22–24] and references therein.

New trends concern the theory of mean-field games with a major player, see [9], the numeric analysis, see [25, 26], and the games with a discrete state space, see [27] and references therein.

Even more recent development deals with mean-field games with common noise, which are only starting to be analyzed. Of course, common noise can be considered as a kind of neutral major player, but the usual setting for the latter [9] introduces the corresponding noise into the coefficients of the SDEs of the minor players, rather than adding additional common stochastic differential. One of the ideas (and results) of our contribution is to use the method of stochastic characteristics to link these two models.

Some simple concrete models of mean-field game types with common noise applied to modeling inter bank loans are analyzed in detail in [28]. A model of common noise with constant coefficients is discussed in [29]. Seemingly first serious contributions to the general theory of mean-field games with common noise and related theory of Master Equation (ME) are the papers [12] and [30], which includes well-posedness for the mean-field limiting evolution under certain assumptions (see e.g. [5–7, 31]). However, [12] and [30] work mostly with controlled SDEs, and our approach is rather different, being based on McKean-Vlasov SPDEs. The references on the literature on McKean-Vlasov equation are given in the Sections devoted to this equations.

3   Our strategy and results

Let us introduce the following conditions:

(C1) Functions $\sigma_{ind}(x), \sigma_{com}(x) \in C^3(\mathbf{R})$ for all $x \in \mathbf{R}$ are positive and never approaching zero;

(C2) Function $b(t, x, \mu, u(t, x, \mu))$ is continuous and bounded on $[0, T] \times \mathbf{R} \times \mathcal{M}_1^{sign}(\mathbf{R}) \times U$, $b(t, ., \mu, u(t, ., \mu)) \in C^2(\mathbf{R})$, and

$$\frac{\partial b}{\partial x}(t, x, ., u(t, x, .)) \in C^1(\mathcal{M}_1^{sign}(\mathbf{R}))$$

with bounds uniform with respect to all variables;

(C3) The first and second order variational derivatives of $b(t, x, \mu, u)$ with respect to μ are well defined, bounded and

$$b(t, x, ., u(t, x, .)) \in (C^{2,1\times1} \cap C^{1,2})(\mathcal{M}_1^{sign}(\mathbf{R})).$$

Our main result is the following.

Theorem 3.1. *Let the functions $b, \sigma_{com}, \sigma_{ind}$ satisfy Conditions (C1)–(C3), and let function $V(t, x, \mu)$ be a solution to problem (15),(16). Assume $J(t, x, \mu, u(t, x, \mu))$ and $V(t, x, \mu)$ belong to the space $(C^{2,1\times1} \cap C^{1,2})(\mathcal{M}_1^{sign}(\mathbf{R}))$ as functions of μ, and belong to the space $C^2(\mathbf{R})$ as functions of x and their derivatives with respect to x belong to the space $C^{1,1}(\mathcal{M}_1^{sign}(\mathbf{R}))$. Then the profile of symmetric strategies $\hat{u}_t(x, \mu)$ given by (13) is an ϵ-Nash equilibrium of the N-player game given by (1), (2), with $\epsilon(N) \sim 1/N$ as $N \to \infty$.*

The proof of Theorem 3.1 will be given at the end of this Section. It is based on the results on the McKean-Vlasov (SPDEs) equation developed in the Sections 4-6.

Remark 2. *The assumptions can be weaken in many ways, but some regularity of the control synthesis u (like being Lipschitz in x, μ) is definitely needed for the rather subtle estimate $1/N$.*

Additionally, in preparation to this result, we obtain two other results of independent interest, not linked with any optimization problem, namely the regularity and sensitivity for McKean-Vlasov SPDE, Theorems 4.2 and 5.1, and the $1/N$-rates of convergence for interacting diffusions to the limiting measure-valued diffusion, Theorem 3.2 (often interpreted as the 'propagation of chaos' property). Notice that the convergence itself is a known result (see e.g. [32] or [4]). The well-posedness of the McKean-Vlasov SPDE was shown in [4] in the class of L_2-functions, and for measures in [32], though under an additional monotonicity assumption.

Let us fix some basic notations for the function spaces. For a topological space X, $C(X)$ denotes the Banach space of continuous functions equipped with the sup-norm $\|.\|$. The topology on measures will be always the weak one. If $X = \mathbf{R}^d$, then $C^k(X)$ denotes the Banach space of functions with all derivatives up to order k belonging to $C(X)$, $L_1(X)$ denotes the space of integrable functions, $L_\infty(X)$ the space of bounded measurable functions with the essential supremum as a norm, $H_1^1(X)$ the Sobolev space of integrable functions such that its generalized derivative is also integrable. If X is not indicated explicitly in this notations we mean $X = \mathbf{R}$.

Let $C^{k\times k}(\mathbf{R}^{2d})$ denote the space of functions f on $\mathbf{R}^{2d}$ such that the partial derivatives

$$\frac{\partial^{\alpha+\beta} f}{\partial x^\alpha \partial y^\beta} \quad \text{with multi-index } \alpha, \beta \text{ such that } |\alpha| \le k, |\beta| \le k,$$

belong to $C(\mathbf{R}^{2d})$.

Remark 3. *The space $C^{k\times k}(\mathbf{R}^{2d})$ looks a bit exotic. However, it is very natural for the study of the second order derivatives of nonlinear measure-valued flows. The spaces of this kind also play an important role in the analysis of stochastic flows in H. Kunita [3], though Kunita's spaces are slightly more general as they allow for a linear growth of functions.*

Recall that for a functional $F(\mu)$ on $\mathcal{M}^{sign}(\mathbf{R}^d)$, the variational derivative is defined as

$$\frac{\delta F}{\delta \mu(x)}[\mu] = \frac{d}{dh}|_{h=0} F(\mu + h\delta_x).$$

Derivatives of higher order are defined accordingly. For instance, if $F = F_j$ is given by (3), then

$$\frac{\delta F}{\delta \mu(x)} = k \int \tilde{F}_j(x, x_2, \cdots, x_k) \mu(dx_2) \cdots \mu(dx_k),$$

$$\frac{\delta^2 F}{\delta \mu(x)\delta \mu(y)} = k(k-1) \int \tilde{F}_j(x, y, x_3, \cdots, x_k) \mu(dx_3) \cdots \mu(dx_k).$$

Let $C^k(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))$ denote the space of functionals such that the kth order variational derivatives are well defined and represent continuous functions. It is a Banach space with the norm

$$\|F\|_{C^k(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))} = \sum_{j=0}^{k} \sup_{x_1,\cdots,x_j,\mu \in \mathcal{M}_\lambda^{sign}(\mathbf{R}^d)} \left| \frac{\delta^j F}{\delta \mu(x_1) \cdots \delta \mu(x_j)} \right|.$$

Let $C^{k,l}(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))$ denote the subspace of $C^k(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))$ such that all derivatives up to order k have continuous bounded derivatives up to order l as functions of their spatial variables. It is a Banach space with the norm

$$\|F\|_{C^{k,l}(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))} = \sum_{j=0}^{k} \sup_{\mu \in \mathcal{M}_\lambda^{sign}(\mathbf{R}^d)} \left\| \frac{\delta^j F}{\delta \mu(.) \cdots \delta \mu(.)}[\mu] \right\|_{C^l(\mathbf{R}^{dj})}.$$

Finally, let $C^{2,k \times k}(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))$ be the space of functionals with the norm

$$\|F\|_{C^{2,k \times k}(\mathcal{M}_\lambda^{sign}(\mathbf{R}^d))} = \sup_{\mu \in \mathcal{M}_\lambda^{sign}(\mathbf{R}^d)} \left\| \frac{\delta^2 F}{\delta \mu(.)\delta \mu(.)} \right\|_{C^{k \times k}(\mathbf{R}^{2d})}.$$

These Banach spaces are natural objects for studying sensitivity for nonlinear measure-valued evolutions. As we are interested mostly in probability measures, we shall usually tacitly assume $\lambda = 1$ for these spaces.

As the derivatives of measures are not always measures (say, the derivative of δ_x is δ_x'), to study the derivatives of the nonlinear evolutions one needs the spaces dual to the spaces of smooth functions. Namely, for a generalized function (distribution) ξ on $\mathbf{R}^d$ we say that it belongs to the space $[C^k(\mathbf{R}^d)]'$ if the norm

$$\|\xi\|_{[C^k(\mathbf{R}^d)]'} = \sup_{\phi: \|\phi\|_{C^k(\mathbf{R}^d)} \le 1} |(\xi, \phi)|$$

is finite. For instance,

$$\|\delta_x^{(k)}\|_{[C^k(\mathbf{R}^d)]'} = 1.$$

We shall use these norms mostly for generalized functions that are given by locally integrable functions. In this case the $[C(\mathbf{R}^d)]'$-norm coincides with the L_1 norm. To see why these spaces are handy, let us observe that if we take a spatial derivative of a heat kernel, then its L_1-norm is of order $t^{-1/2}$ for small t, but its $[C'(\mathbf{R}^d)]'$-norm is uniformly bounded.

Let us explain our strategy for proving Theorem 3.1.

For any N and a fixed common strategy $u_t(x, \mu)$, solutions to the system of SDEs (1) on $t \in [0, T]$ define a backward propagator (also referred in the literature as a flow or as a two-parameter semigroup) $U_N^{s,t} = U_N^{s,t}[u(.)]$, $0 \leq s \leq t \leq T$, of linear contractions on the space $C_{sym}(\mathbf{R}^N)$ of symmetric functions via the formula

$$(U_N^{s,t} f)(x_1, \cdots, x_N) = \mathbf{E}f(X_1, \cdots, X_N)_{s,t}(x_1, \cdots, x_N), \tag{19}$$

where $(X_1, \cdots, X_N)_{s,t}(x_1, \cdots, x_N)$ is the solution to (1) at time t with the initial condition

$$(X_1, \cdots, X_N)_{s,s}(x_1, \cdots, x_N) = (x_1, \cdots, x_N)$$

at time s. The corresponding dual forward propagator $V_N^{t,s} = (U_N^{s,t})'$ is defined by the equation

$$(f, V_N^{t,s} \mu) = (U_N^{s,t} f, \mu). \tag{20}$$

It acts on the probability measures on $\mathbf{R}^N$, so that if μ is the initial distribution of $(X_1, \cdots, X_N)$ at time s, then $V_N^{t,s} \mu$ is the distribution of $(X_1, \cdots, X_N)$ at time t.

By the standard inclusion

$$(x_1, \cdots, x_N) \to \frac{1}{N}(\delta_{x_1} + \cdots + \delta_{X_N}) \tag{21}$$

the set $\mathbf{R}^N$ is mapped to the set $\mathcal{P}_N(\mathbf{R})$ of normalized sums of N Dirac's measures, so that $U_N^{s,t}$, $V_N^{t,s}$ can be considered as propagators in $C(\mathcal{P}_N(\mathbf{R}))$ and $\mathcal{P}(\mathcal{P}_N(\mathbf{R}))$ respectively.

On the other hand, for a fixed function $u_t(x, \mu)$, the solution of SPDE (5) specifies a stochastic process, a diffusion, on the space of probability measures $\mathcal{P}(\mathbf{R})$ defining the backward propagator $U^{s,t} = U^{s,t}[u(.)]$ on $C(\mathcal{P}(\mathbf{R}))$:

$$(U^{s,t} f)(\mu) = \mathbf{E}f(\mu_{s,t}(\mu)), \tag{22}$$

where $\mu_{s,t}(\mu)$ is the solution to (5) at time t with a given initial condition μ at time $s \leq t$.

From the convergence of the empirical measures μ_t^N, mentioned above, it follows that $U_N^{s,t}$ tend $U^{s,t}$, as $N \to \infty$. The following result provides the rates for the weak convergence.

Theorem 3.2. *Let the functions $b, \sigma_{com}, \sigma_{ind}$ satisfy Conditions (C1)–(C3). Then, for any $\mu \in \mathcal{P}_N(\mathbf{R})$ and $F \in (C^{2,1\times1} \cap C^{1,2})(\mathcal{M}_1^{sign}(\mathbf{R}))$*

$$\|(U^{s,t} - U_N^{s,t})F(\mu)\|_{C(\mathcal{M}_1^{sign}(\mathbf{R}))} \leq \frac{C(T)}{N} \left(\|F\|_{C^{2,1\times1}(\mathcal{M}_1^{sign}(\mathbf{R}))} + \|F\|_{C^{1,2}(\mathcal{M}_1^{sign}(\mathbf{R}))} \right) \tag{23}$$

for $0 \leq s \leq t \leq T$.

Proof. Let us return to our initial equation (1). By the standard assumption of the Lipschitz continuity of all coefficients, equation (1) is well-posed in $\mathbf{R}^N$ and specifies a Feller diffusion and the corresponding backward and forward propagators U_N, V_N given by (19), (20). We are interested in the limit of this diffusion as $N \to \infty$.

Applying Ito's formula we obtain the generator of the diffusion specified by (1):

$$A_N f(x_1, \cdots, x_N) = \sum_{j=1}^{N} B_\mu^i f + \sum_{i<j} \sigma_{com}(x_i)\sigma_{com}(x_j)\frac{\partial^2 f}{\partial x_i \partial x_j}, \tag{24}$$

where $\mu = (\delta_{x_1} + \cdots + \delta_{x_N})/N$ and

$$B_\mu g(x) = b(t, x, \mu, u(t, x, \mu))\frac{\partial g}{\partial x} + \frac{1}{2}(\sigma_{com}^2(x) + \sigma_{ind}^2(x))\frac{\partial^2 g}{\partial x^2}, \tag{25}$$

with B_μ^i denoting the action of B_μ on the ith coordinate of f.

Here and everywhere by a time-dependent generator, say A_N above, of a non-homogeneous Markov process we mean a time-dependent family of operators such that for f from some invariant dense subspace of bounded continuous functions the equation

$$\frac{d}{ds}U_N^{s,t}f = -A_N U_N^{s,t}f$$

holds for $s \le t$. In the case of the N-particle diffusion, the invariant subspace can be usually taken to be the space of twice differentiable functions (which are invariant if σ and b are twice and once differentiable respectively). In the case of the limiting measure-valued process the invariant domains will be given by the subspaces $C^{2,k\times k}(\mathcal{M}^{sign})$.

The first term in (24) can be considered as describing a diffusion arising from the system of particles with a mean-field interaction and the second term as giving an additional binary interaction (though not of a standard potential type that can be easily included in the mean-field interaction).

By the standard inclusion (21), the process specified by (24) can be equivalently considered as a measure-valued process defined on the set of linear combinations $\mathcal{P}_N(\mathbf{R})$ of the Dirac atomic measures. On the level of propagators this correspondence arises from the identification of symmetric functions f on X^N with the functionals $F = F_f$ on $\mathcal{P}_N(\mathbf{R})$ via the equation

$$f(x_1, \cdots, x_N) = F_f[(\delta_{x_1} + \cdots + \delta_{x_N})/N].$$

To recalculate the generator (24) in terms of functionals F on measures we use the following simple formulas for differentiation of functionals on measures (proofs can be found e.g. in [33]): for $\mu = h(\delta_{x_1} + \cdots + \delta_{x_N})$ with $h = 1/N$

$$\frac{\partial}{\partial x_j}F(\mu) = h\frac{\partial}{\partial x_j}\frac{\delta F(\mu)}{\delta\mu(x_j)}, \tag{26}$$

$$\frac{\partial^2}{\partial x_j^2}F(\mu) = h\frac{\partial^2}{\partial x_j^2}\frac{\delta F(\mu)}{\delta\mu(x_j)} + h^2\frac{\partial^2}{\partial y\partial z}\frac{\delta^2 F(\mu)}{\delta\mu(y)\delta\mu(z)}\Big|_{y=z=x_j}, \tag{27}$$

$$\frac{\partial^2}{\partial x_i\partial x_j}F(\mu) = h^2\frac{\partial^2}{\partial x_i\partial x_j}\frac{\delta^2 F(\mu)}{\delta\mu(x_i)\delta\mu(x_j)}, \quad i \ne j. \tag{28}$$

Applying these formulas in conjunction with the obvious identity

$$h^2\sum_{i<j:i,j\in\{1,\dots,N\}}\phi(x_i,x_j) = \frac{1}{2}\int\int\phi(z_1,z_2)\mu(dz_1)\mu(dz_2) - \frac{h}{2}\int\phi(z,z)\mu(dz), \tag{29}$$

leads to the following expression of A_N in terms of $F(\mu)$ (for details and more general calculations see [33]):

$$A_N F(\mu) = \Lambda_{lim}F(\mu) + \frac{1}{N}\Lambda_{corr}F(\mu), \tag{30}$$

with

$$\Lambda_{lim}F(\mu) = \int_{\mathbf{R}}\left(B_\mu\frac{\delta F}{\delta\mu(.)}\right)(y)\mu(dy)$$
$$+\frac{1}{2}\int_{\mathbf{R}^2}\sigma_{com}(y)\sigma_{com}(z)\frac{\partial^2}{\partial y\partial z}\frac{\delta^2 F}{\delta\mu(y)\delta\mu(z)}\mu(dy)\mu(dz), \tag{31}$$

$$\Lambda_{corr}F(\mu) = \frac{1}{2}\int_{\mathbf{R}}\sigma_{ind}^2(x)\left.\frac{\partial^2}{\partial y\partial z}\frac{\delta^2 F(\mu)}{\delta\mu(y)\delta\mu(z)}\right|_{y=z=x}\mu(dx). \tag{32}$$

Thus we have an explicit expression for the limit of A_N as $N \to \infty$ and for the correction term, which are well defined for functional F from the spaces $C^{1,2}(\mathcal{M}^{sign}) \cap C^{2,1\times1}(\mathcal{M}^{sign})$.

It is straightforward to check by Ito's formula that the operator Λ_{lim} generates the measure-valued process defined by the solution of equation (1). Hence we have the convergence of the generators of N-particle approximations to the generator of the process given by (1) on the space $C^{1,2}(\mathcal{M}^{sign}) \cap C^{2,1\times1}(\mathcal{M}^{sign})$ with the uniform rate of convergence of order $1/N$.

But according to Theorem 6.1 (see below), the propagator of the process generated by (1) acts by bounded operators on this subspace. Hence Theorem 3.2 follows from the standard representation of the difference of two propagators in terms of the difference of their generators:

$$U_N^{t,r} - U^{t,r} = \int_t^r U_N^{t,s}(A_N - \Lambda_{lim})_s U^{s,r}ds. \tag{33}$$

$\square$

This result belongs to the statistical mechanics of interacting diffusions (see a detail discussion and bibliography in [33]), so that its significance goes beyond any links with games or control theory.

This result is not sufficient for us, as we have to allow one of the agent to behave differently from the others. To tackle this case we shall considered the corresponding problem with a tagged agent. Namely, consider the Markov process on pairs $(X_t^{1,N}, \mu_t^N)[u^{ind}(.), u^{com}(.)]$, where u^{ind} and u^{com} are some U-valued functions $u_t^{ind}(x,\mu)$, $u_t^{com}(x,\mu)$, $(X_t^{1,N}, \cdots, X_t^{N,N})$ solves (1) under the assumptions that the first agent uses the control $u_t^{ind}(X_t^{1,N}, \mu_t^N)$ and all other agents $i \neq 1$ use the control $u_t^{com}(X_t^{i,N}, \mu_t^N)$, and $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$.

Remark 4. *The coordinates $(X_t^{1,N}, \mu_t^N)$ of our pair process are not independent. Quite opposite, $X_t^{1,N}$ is the position of the first δ-function in μ_t^N. However, we are aiming at the limit $N \to \infty$ where the influence of $X_t^{1,N}$ on μ_t^N becomes negligible, and we do not want it to be lost at the limit. Alternatively, to avoid this dependence, one can consider (as some authors do), instead of our μ_t^N, the measures that do not take $X_t^{1,N}$ into account, that is $\tilde{\mu}_t^N = \frac{1}{N}\sum_{i=2}^N \delta_{X_t^{i,N}}$, but this would neither change the results, nor simplifies the notations.*

Let us now define the corresponding tagged propagators $U_{N,tag}^{s,t} = U_{N,tag}^{s,t}[u^{ind}(.), u^{com}(.)]$ and $U_{tag}^{s,t} = U_{tag}^{s,t}[u^{ind}(.), u^{com}(.)]$:

$$(U_{N,tag}^{s,t}F)(x,\mu) = \mathbf{E}F(X_t^{1,N}, \mu_t^N)[u^{ind}(.), u^{com}(.)](x,\mu), \tag{34}$$

11

where $\mu = \frac{1}{N}\sum_{j=1}^{N}\delta_{x_j}$ is the position of the process at time s and where $x = x_1$;

$$(U_{tag}^{s,t}F)(x,\mu) = \mathbf{E}F(X_t^1,\mu_t)[u^{ind}(.),u^{com}(.)](x,\mu), \tag{35}$$

where the process $(X_t^1,\mu_t)[u^{ind}(.),u^{com}(.)](x,\mu)$ with the initial data x,μ at time s is the solution to the system of stochastic equations

$$dX_t^1 = b(t,X_t^1,\mu_t,u_t^{ind}(X_t^1,\mu_t)) + \sigma_{ind}(X_t^1)dB_t^1 + \sigma_{com}(X_t^1)dW_t, \tag{36}$$

$$d(\phi,\mu_t) = (L[t,\mu_t,u_t^{com}(.,\mu_t)]\phi,\mu_t)\,dt + (\sigma_{com}(.)\nabla\phi,\mu_t)\,dW_t \tag{37}$$

(the second equation is actually independent of the first one).

The following is the basic convergence result for the tagged processes.

Theorem 3.3. *Under the assumptions of Theorem 3.2 (with both u_t^{com},u_t^{ind} satisfying these assumptions), let $F(x,\mu)$, $x\in\mathbf{R}$, belongs to the space $(C^{2,1\times1}\cap C^{1,2})(\mathcal{M}_1^{sign}(\mathbf{R}))$ as a function of μ, $F\in C^2(\mathbf{R})$ as a function of x and $\frac{\partial F}{\partial x}(x,.)\in C^{1,1}(\mathcal{M}_1^{sign}(\mathbf{R}))$. Then, for any $\mu\in\mathcal{P}_N(\mathbf{R})$*

$$\|(U_{tag}^{s,t} - U_{N,tag}^{s,t})F\|_{C(\mathbf{R}\times\mathcal{M}_1^{sign}(\mathbf{R}))}$$

$$\leq \frac{C(T)}{N}\left(\sup_x\|F(x,.)\|_{C^{2,1\times1}(\mathcal{M}_1^{sign}(\mathbf{R}))} + \sup_x\|F(x,.)\|_{C^{1,2}(\mathcal{M}_1^{sign}(\mathbf{R}))}\right.$$

$$\left.+ \sup_\mu\|F(.,\mu)\|_{C^2(\mathbf{R})} + \sup_x\|\frac{\partial F}{\partial x}(x,.)\|_{C^{1,1}(\mathcal{M}_1^{sign}(\mathbf{R}))}\right). \tag{38}$$

Proof. The well-posedness of the process on pairs (x,μ) solving equations (36) and (37) is straightforward once the well-poesdness of the process solving (37) is proved, because equation (37) does not depend on x, and once it is solved, equation (36) is just a usual Ito's equation. Straightforward extension of the above calculations for the generator of the process solving (37) show that the process solving (36) - (37) is generated by the operator

$$\Lambda_{lim}F(x,\mu) + \tilde{\Lambda}_{Lim}F(x,\mu),$$

where Λ_{lim} is given by (31) and acts on the variable μ,

$$\tilde{\Lambda}_{lim}F(x,\mu) = b(t,x,\mu,u_t^{ind}(x,\mu))\frac{\partial F}{\partial x} + \frac{1}{2}(\sigma_{ind}^2 + \sigma_{com}^2)(x)\frac{\partial^2 F}{\partial x^2}$$

$$+ \int\sigma_{com}(x)\sigma_{com}(y)\frac{\partial^2}{\partial x\partial y}\frac{\delta F}{\delta\mu(y)}\mu(dy), \tag{39}$$

and with the same correction term (32). Thus the proof of Theorem 3.3 is the same as for Theorem 3.2. $\qquad\square$

Proof of Theorem 3.1. Let u_1 be any adaptive control of the first player and V_1 the corresponding payoff in the game of N players, where all other players are using $u_{com}(t,x,\mu)$ arising from a solution to (15),(16). Then $V_1 \geq V_2$, where V_2 is obtained by playing optimally, that is using control u_2 arising from the solution to (11). By Theorem 3.3,

$$|V_2 - V_{2,lim}| \leq C/N,$$

where $V_{2,lim}$ is obtained by playing u_2 in the limiting game specified by equations (36), (37). But $V_{2,lim} \geq V$, where V is the optimal payoff for the first player in the limiting game of two players, where the second, measure-valued, player uses u_{com}. Consequently,

$$V_1 \geq V_2 \geq V_{2,lim} - \frac{C}{N} \geq V - \frac{C}{N},$$

completing the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

4 On the regularity of McKean-Vlasov SPDEs

In this and the next sections we develop the sensitivity analysis for McKean-Vlasov SPDEs, which, on the one hand side, represent an important ingredient in the proof of our main result on mean-field games, but on the other hand, has an independent significance for the theory of SPDEs. Notice that there is quite an extensive literature on the properties of equation (40) with $A = 0$ (see e.g. [34], [35] and references therein), but for $A \neq 0$ much less is known, so that even the regularity results from Theorem 4.2 below seem to be new.

For a function $v(t, x)$, $t \geq 0, x \in \mathbf{R}$, let us consider the stochastic equation

$$dv = L_t(v)\, dt + \Omega v \circ dW_t, \tag{40}$$

where W_t is a one-dimensional Brownian motion,

$$\Omega v(x) = A(x)\frac{\partial v}{\partial x} + B(x)v(x), \tag{41}$$

$$L_t(v) = \frac{1}{2}\sigma^2(x)\frac{\partial^2 v}{\partial x^2} + b(t, x, [v])\frac{\partial v}{\partial x} + c(t, x, [v])v, \tag{42}$$

with some functions $A(x), B(x), \sigma(x)$ and the functions b, c depending in a smooth way on the function (or a measure) v. To visualize this dependence, one can think of b, c depending on v via a finite set of moments of type

$$F_j[v] = \int \tilde{F}_j(x_1, \cdots, x_k)v(x_1) \cdots v(x_k)\, dx_1 \cdots dx_k, \tag{43}$$

with some bounded symmetric measurable functions $\tilde{F}_j$.

In (40), $\circ$ denotes the Stratonovich differential. From the usual rule $Y \circ dX = Y\, dX + \frac{1}{2}dY\, dX$, one can rewrite (40) as an equation with Ito's differential of the similar kind:

$$dv = L_t(v)\, dt + \Omega v\, dW_t + \frac{1}{2}\Omega^2 v\, dt, \tag{44}$$

or explicitly

$$dv = L_t(v)\, dt + \Omega v\, dW_t + \frac{1}{2}\left[A^2\frac{\partial^2 v}{\partial x^2} + A(2B + A')\frac{\partial v}{\partial x} + (AB' + B^2)v\right]dt. \tag{45}$$

Our objective is to study the well-posedness of equation (40) and more importantly its sensitivity with respect to initial conditions.

Our main assumptions will be that

$$0 < \sigma_1 \leq \sigma(x) \leq \sigma_2, \quad 0 < A_1 \leq A(x) \leq A_2, \tag{46}$$

that $\sigma \in C^2(\mathbf{R}), B \in C^2(\mathbf{R}), A \in C^3(\mathbf{R})$ so that

$$\sup_x \max(|\sigma'(x)|, |\sigma''(x)|, |B(x)|, |B'(x)|, |B''(x)|, |A'(x)|, |A''(x)|, |A'''(x)|) \leq B_1, \tag{47}$$

and

$$\max \left(\|b(t, ., [v])\|_{C^1(\mathbf{R})}, \|c(t, ., [v])\|_{C(\mathbf{R})} \right) \leq b_1, \tag{48}$$

$$\sup_{t,y} \sup_{\|v\|_{\mathcal{M}(\mathbf{R})} \leq \lambda} \max \left(\left\| \frac{\delta b(t, y, [v])}{\delta v(.)} \right\|_{C^1(\mathbf{R})}, \left\| \frac{\delta c(t, y, [v])}{\delta v(.)} \right\|_{C(\mathbf{R})}, \left\| \frac{\delta}{\delta v(.)} \frac{\partial b(t, y, [v])}{\partial y} \right\|_{C(\mathbf{R})} \right) \leq C(\lambda), \tag{49}$$

with some constants $\sigma_1, \sigma_2, A_1, A_2, B_1, b_1$ and a function $C(\lambda)$.

As above, we shall often omit the arguments of various functions. Moreover sometimes, we shall write $v(t, x)$ as $v_t(x)$ when stressing that certain operator acts on v as a function of x for a given t.

Our basic approach will be the method of stochastic characteristics, see [3], [36], though in its simplest form, available for one-dimensional noise. This method allows one to turn equation (40) into a non-stochastic equation of the second order, but with random coefficients. Namely, for $A(x), B(x) \in C^1(\mathbf{R})$, operator (41) generates a contraction group $e^{t\Omega}$ in $C(\mathbf{R})$, so that $e^{t\Omega}v_0(x)$ is the unique solution to the equation

$$\frac{\partial v}{\partial t} = \Omega v$$

with the initial condition $v(0, x) = v_0(x)$. Explicitly,

$$e^{t\Omega}v_0(x) = v_0(Y(t, x))G(t, x), \quad t \in \mathbf{R}, \tag{50}$$

where $Y(t, x)$ is the unique solution to the ODE $\dot{Y} = -A(Y)$ with the initial condition $Y(0, x) = x$ and

$$G(t, x) = \exp\{\int_0^t B(Y(s, x))\, ds\}.$$

In particular, G has the properties:

$$G(-t, x) = G^{-1}(t, Y(-t, x)) = \exp\{\int_0^{-t} B(Y(s, x))\, ds\} = \exp\{-\int_{-t}^0 B(Y(s, x))\, ds\},$$

$$\frac{1}{G}\frac{\partial G}{\partial x}(t, x) = \frac{\partial \ln G}{\partial x}(t, x) = \int_0^t B'(Y(s, x))\frac{\partial Y}{\partial x}(s, x)\, ds.$$

Since the product-rule of calculus is valid for the Stratonovich differentials, making the change of unknown function v to $g = \exp\{-\Omega W_t\}v$ rewrites (40) in terms of g as

$$\dot{g}_t = \tilde{L}_t[W](g_t) = \exp\{-\Omega W_t\}L_t(\exp\{\Omega W_t\}g_t), \tag{51}$$

with $\dot{v}$ denoting the usual derivative of a function v in time t. Of course one can obtain the same result using usual Ito's formula and equation (44). Since the operators $e^{t\Omega}$ form

a bounded semigroup in $L_1(\mathbf{R})$, as well as in $C^k(\mathbf{R})$ and $C^k_\infty(\mathbf{R})$ whenever $A, B \in C^k(\mathbf{R})$, equations (51) and (40) are equivalent in the strongest possible sense.

To have a more concrete version of (51) we calculate

$$\frac{\partial}{\partial x}(\exp\{\Omega W_t\}g_t)(x) = \frac{\partial G}{\partial x}(W_t, x)g_t(Y(W_t, x)) + G(W_t, x)\frac{\partial g_t}{\partial Y}(Y(W_t, x))\frac{\partial Y}{\partial x}(W_t, x),$$

$$\frac{\partial^2}{\partial x^2}(\exp\{\Omega W_t\}g_t)(x) = \frac{\partial^2 G}{\partial x^2}(W_t, x)g_t(Y(W_t, x)) + 2\frac{\partial G}{\partial x}(W_t, x)\frac{\partial g_t}{\partial Y}(Y(W_t, x))\frac{\partial Y}{\partial x}(W_t, x)$$

$$+ G(W_t, x)\frac{\partial g_t}{\partial Y}(Y(W_t, x))\frac{\partial^2 Y}{\partial x^2}(W_t, x) + G(W_t, x)\frac{\partial^2 g_t}{\partial Y^2}(Y(W_t, x))\left(\frac{\partial Y}{\partial x}(W_t, x)\right)^2.$$

Hence,

$$\tilde{L}_t[W](g_t)(x) = \frac{1}{2}\tilde{\sigma}^2(x)\frac{\partial^2 g_t}{\partial x^2} + \tilde{b}(t, x, [g_t])\frac{\partial g_t}{\partial x} + \tilde{c}(t, x, [g_t])g_t, \tag{52}$$

with

$$\tilde{\sigma}^2(x) = \sigma^2(Y(-W_t, x))\left(\frac{\partial Y}{\partial z}(W_t, z)|_{z=Y(-W_t, x)}\right)^2, \tag{53}$$

$$\tilde{b}(t, x, [g]) = \left(b(t, z, [\exp\{\Omega W_t\}g])\frac{\partial Y}{\partial z}(W_t, z)\right)|_{z=Y(-W_t, x)}$$

$$+ \left[\frac{1}{2}\sigma^2(z)\left(\frac{\partial^2 Y}{\partial z^2}(W_t, z) + 2\frac{\partial \ln G}{\partial z}(W_t, z)\frac{\partial Y}{\partial z}(W_t, z)\right)\right]|_{z=Y(-W_t, x)}, \tag{54}$$

$$\tilde{c}(t, x, [g]) = \left(c(t, z, [\exp\{\Omega W_t\}g]) + b(t, z, [\exp\{\Omega W_t\}g])\frac{\partial \ln G}{\partial z}(W_t, z)\right)\Big|_{z=Y(-W_t, x)}$$

$$+ \left[\frac{1}{2}\sigma^2(z)\left(\frac{1}{G}\frac{\partial^2 G}{\partial z^2}\right)(W_t, z)\right]\Big|_{z=Y(-W_t, x)}. \tag{55}$$

The formulas above have straightforward extension to x from arbitrary dimension. The simplification arising from working in one-dimension is as follows:

$$Y(t, x) = \Phi^{-1}(t + \Phi(x)),$$

where

$$\Phi(y) = \int_0^y \frac{dz}{A(z)}.$$

Hence, under (46), (47) it follows that

$$\frac{A_1}{A_2} \le \frac{\partial Y}{\partial x}(t, x) \le \frac{A_2}{A_1}, \quad \left|\frac{\partial^2 Y}{\partial x^2}(t, x)\right| \le \frac{B_1 A_2}{A_1^2}\left(1 + \frac{A_2^2}{A_1^2}\right)$$

and

$$\left|\frac{\partial \ln G}{\partial x}(t, x)\right| \le |t|B_1\frac{A_2}{A_1}$$

for all t, x and hence

$$\sigma_1\frac{A_1}{A_2} \le \tilde{\sigma}(x) \le \sigma_2\frac{A_2}{A_1}, \tag{56}$$

$$\tilde{b}(t, x, [g]) \le C(T)(1 + \bar{W}_T), \quad \tilde{c}(t, x, [g]) \le C(T)(1 + \bar{W}_t), \tag{57}$$

with some constants $C(T)$ and $\bar{W}_T = \max_{t \in [0,T]} |W_t|$.

Thus on any finite interval of time $[0, T]$ equation

$$\dot{g}_t = \tilde{L}_t[W]g_t \tag{58}$$

is the usual nonlinear McKean-Vlasov diffusion equation with uniformly elliptic second order part and bounded coefficients. For this equation both well-posedness and smooth dependence on initial condition is known, see e.g. [33] and [37]. A new point for us is the necessity to have bounds for the expectations of the various relevant objects. So we shall briefly recall the argument used for the analysis of the sensitivity of equation (58) paying attention to the latter issue.

Let us first make a precise statement about equation (58) independently on its link with our initial SPDE. We shall need the following assumptions:

$$\tilde{\sigma}_1 \leq \tilde{\sigma} \leq \tilde{\sigma}_2, \quad \tilde{\sigma}' \leq \tilde{\sigma}_2, \tag{59}$$

$$\max \left(\|\tilde{b}(t, ., [v])\|_{C^1(\mathbf{R})}, \|\tilde{c}(t, ., [v])\|_{C(\mathbf{R})} \right) \leq \tilde{b}(1 + \bar{W}_T) \tag{60}$$

and either

$$\sup_{\|v\|_{L_1} \leq \lambda} \max \left(\left\| \frac{\delta \tilde{b}(t, y, [v])}{\delta v(.)} \right\|_{C(\mathbf{R})}, \left\| \frac{\delta \tilde{c}(t, y, [v])}{\delta v(.)} \right\|_{C(\mathbf{R})}, \left\| \frac{\delta}{\delta v(.)} \frac{\partial \tilde{b}(t, y, [v])}{\partial y} \right\|_{C(\mathbf{R})} \right) \leq \tilde{C}(\bar{W}_T, \lambda) \tag{61}$$

or

$$\sup_{\|v\|_{C(\mathbf{R})} \leq \lambda} \max \left(\left\| \frac{\delta \tilde{b}(t, y, [v])}{\delta v(.)} \right\|_{L_1(\mathbf{R})}, \left\| \frac{\delta \tilde{c}(t, y, [v])}{\delta v(.)} \right\|_{L_1(\mathbf{R})}, \left\| \frac{\delta}{\delta v(.)} \frac{\partial \tilde{b}(t, y, [v])}{\partial y} \right\|_{L_1(\mathbf{R})} \right) \leq \tilde{C}(\bar{W}_T, \lambda) \tag{62}$$

for some constants $\tilde{\sigma}_{1,2}, \tilde{b}$ and a function $\tilde{C}$ depending on $\bar{W}_T = \sup_{t \in [0,T]} |W_t|$ and λ.

The idea is to rewrite the Cauchy problem for (58) with the initial condition g_0 in the mild form, that is as a fixed point equation

$$\Phi[g] = g \tag{63}$$

for the mapping

$$\Phi g \mapsto \Phi_t[g](x) = \int G(t, x, y) g_0(y) dy$$

$$+ \int_0^t \int G(t - s, x, y) \left[\tilde{b}(s, y, [g_s]) \frac{\partial g_s}{\partial y}(y) + \tilde{c}(s, y, [g_s]) g_s(y) \right] dy\, ds, \tag{64}$$

where $G(t, x, y)$ is the Green function for the Cauchy problem of the equation

$$\dot{g} = \frac{1}{2} \tilde{\sigma}^2(x) \frac{\partial^2 g}{\partial x^2}.$$

Using integration by parts Φ rewrites equivalently as

$$\Phi_t[g](x) = \int G(t, x, y) g_0(y) dy$$

16

$$+ \int_0^t ds \int \left[G(t-s,x,y)\tilde{c}(s,y,[g_s]) - \frac{\partial}{\partial y}(G(t-s,x,y)\tilde{b}(s,y,[g_s])) \right] g_s(y)\, dy, \quad (65)$$

from which it is seen that Φ maps bounded families of functions g_t, $t \in [0,T]$, with a given g_0, to itself, where 'bounded' can be understood either in sup-norm or in L_1-norm.

The Green function G is random, i.e. it depends on W. However, by the standard theory of the second order equations (see [38] and [39]), assuming (59), the function G has the two-sided Gaussian bounds

$$\frac{C_1}{\sqrt{2\pi t}} \exp\{-\frac{(x-y)^2}{2tC_3}\} \le G(t,x,y) \le \frac{C_2}{\sqrt{2\pi t}} \exp\{-\frac{(x-y)^2}{2tC_4}\},$$

and the bound for the derivatives

$$\max\left(\left| \frac{\partial G(t,x,y)}{\partial x} \right|, \left| \frac{\partial G(t,x,y)}{\partial y} \right| \right)$$

$$\le C_5 \left| \frac{\partial}{\partial x} \frac{C_2}{\sqrt{2\pi t}} \exp\{-\frac{(x-y)^2}{2tC_4}\} \right| \le \frac{C_5 C_2}{C_4 \sqrt{t}} \frac{|x-y|}{\sqrt{t}\sqrt{2\pi t}} \exp\{-\frac{(x-y)^2}{2tC_4}\},$$

with constants $C_1 - C_5$ independent of the noise.

This allows one to infer the estimates

$$\|\Phi_t[g^1] - \Phi_t[g^2]\|_{L_1} \le \int_0^t ds \frac{1}{\sqrt{t-s}} \|g_s^1 - g_s^2\|_{L_1} \tilde{C}(\bar{W}_T, \sup_{s \in [0,T]} \max(\|g_s^1\|_{L_1}, \|g_s^2\|_{L_1})) \quad (66)$$

in case of conditions (59), (60), (61) or

$$\|\Phi_t[g^1] - \Phi_t[g^2]\|_{C(\mathbf{R})}$$

$$\le \int_0^t ds \frac{1}{\sqrt{t-s}} \|g_s^1 - g_s^2\|_{C(\mathbf{R})} \tilde{C}(\bar{W}_T, \sup_{s \in [0,T]} \max(\|g_s^1\|_{C(\mathbf{R})}, \|g_s^2\|_{C(\mathbf{R})})) \quad (67)$$

in case of conditions (59), (60), (62).

From these estimates one can infer the convergence of the iterates $\Phi^n(g_0)$ in either sup-norm or L_1-norm and hence the existence of the unique solution $g_t \in L_1$ for any initial $g_0 \in L_1$ (and even for any initial finite measure g_0) or of the unique solution $g_t \in C(\mathbf{R})$ for any initial $g_0 \in C(\mathbf{R})$ (and even for any initial bounded measurable g_0), whenever one can prove the uniform boundedness of the norms of all iterations.

Let us see how one can get an estimate for the norm of the iterations. From the definition of Φ and the estimates of the Green function G we get

$$\|\Phi_t^n\| \le C\|g_0\| + C\int_0^t (t-s)^{-1/2} \|\Phi_s^{n-1}\| ds,$$

where $C = C(T)(1+\bar{W}_T)$ with $C(T)$ a non-random constant, and where $\|\Phi_t^n\|$ is the norm of the nth iteration of Φ applied initially on g_0, and the norm is either in $C(\mathbf{R})$ or in L_1. From this we deduce, by a straightforward induction, that

$$\|\Phi_t^n\| \le C\|g_0\|(1 + C\sqrt{\pi} I_{1/2}\mathbf{1}(t) + \cdots + (C\sqrt{\pi})^{n-1} I_{(n-1)/2}\mathbf{1}(t) + C^{n-1}\pi^{n/2} I_{n/2}\mathbf{1}(t)), \quad (68)$$

were $I_{n/2}\mathbf{1}(t)$ is the application of the fractional integral of order $n/2$ to the constant function $\mathbf{1}$ (that equals one). And consequently we get for the limiting norm of the fixed point the bound in terms of a Mittag-Leffler function and hence eventually in terms of an exponent of Ct. Hence, since the expectation of $\exp\{\bar{W}_T\}$ is finite, we can deduce the bound for the expectation of the fixed point yielding the following result.

Theorem 4.1. *(i) Under assumptions* (59), (60), (61) *any $T > 0$ for any $g_0 \in \mathcal{M}(\mathbf{R})$ there exists a unique solution g_t of equation* (40) *on $[0, T]$ such that $g_t \in L_1(\mathbf{R})$ for all $t > 0$, positive whenever g_0 is positive, and*

$$\|g_t\|_{L_1} \leq C_1(T) \exp\{C_1(T)\bar{W}_T\}\|g_0\|_{\mathcal{M}(\mathbf{R})}, \quad \mathbf{E}\|g_t\|_{L_1} \leq C_2(T)\|g_0\|_{\mathcal{M}(\mathbf{R})} \qquad (69)$$

with constants $C_{1,2}(T)$.

Moreover, if $u_0 \in H_1^1$, then

$$\|g_t\|_{H_1^1} \leq C_1(T) \exp\{C_1(T)\bar{W}_T\}\|g_0\|_{H_1^1}, \quad \mathbf{E}\|g_t\|_{H_1^1} \leq C_2(T)\|g_0\|_{H_1^1}. \qquad (70)$$

Finally, for any $g_0 \in \mathcal{M}(\mathbf{R})$, $g_t \in H_1^1$ a.s. for all $t > 0$ and, if the bounds on the r.h.s. of (61) *do not depend on $\bar{W}_T$, one has the estimate (uniform with respect to the noise)*

$$\|g_t\|_{H_1^1} \leq C_3(T)\frac{1}{\sqrt{t}}\|g_0\|_{\mathcal{M}(\mathbf{R})}. \qquad (71)$$

(ii) Under assumptions (59), (60), (62), *for any $g_0 \in L_\infty(\mathbf{R})$ there exists a unique solution g_t of equation* (40) *on $[0, T]$ such that $g_t \in C(\mathbf{R})$ for all $t > 0$, and*

$$\|g_t\|_{C(\mathbf{R})} \leq C_1(T) \exp\{C_1(T)\bar{W}_T\}\|g_0\|_{L_\infty}, \quad \mathbf{E}\|g_t\|_{C(\mathbf{R})} \leq C_2(T)\|g_0\|_{L_\infty} \qquad (72)$$

with constants $C_{1,2}(T)$.

Moreover, if $g_0 \in C^1(\mathbf{R})$, then

$$\|g_t\|_{C^1(\mathbf{R})} \leq C_1(T) \exp\{C_1(T)\bar{W}_T\}\|g_0\|_{C^1(\mathbf{R})}, \quad \mathbf{E}\|g_t\|_{C^1(\mathbf{R})} \leq C_2(T)\|g_0\|_{C^1(\mathbf{R})}. \qquad (73)$$

Finally, for any $g_0 \in L_\infty(\mathbf{R})$, $g_t \in C^1(\mathbf{R})$ a.s. for all $t > 0$ and, if the bounds on the r.h.s. of (62) *do not depend on $\bar{W}_T$, one has the estimate*

$$\|g_t\|_{C^1(\mathbf{R})} \leq C_3(T)\frac{1}{\sqrt{t}}\|g_0\|_{C(\mathbf{R})}. \qquad (74)$$

Proof. Let us talk about (i) only, as (ii) is fully analogous. The proof of the first statement was already sketched above. The estimates for the norm in H_1^1 are obtained from the iterations in a fully analogous way leading to (70). Finally, we get from (63) the estimate

$$\|g_t\|_{H_1^1} \leq \frac{1}{\sqrt{t}}\|g_0\|_{L_1} + \int_0^t C(\bar{W}_T)\frac{1}{\sqrt{t-s}}\|g_s\|_{H_1^1}ds,$$

so that

$$\sup_{s\in(0,t]}(\sqrt{s}\|g_s\|_{H_1^1}) \leq \|g_0\|_{L_1} + \sqrt{t}C(\bar{W}_T)\int_0^t \frac{1}{\sqrt{t-s}\sqrt{s}}\sup_{s\in(0,t]}(\sqrt{s}\|g_s\|_{H_1^1})\,ds,$$

and thus

$$\sup_{s\in(0,t]}(\sqrt{s}\|g_s\|_{H_1^1}) \leq \|g_0\|_{L_1} + \sqrt{t}C(\bar{W}_T)\int_0^1 \frac{du}{\sqrt{1-u}\sqrt{u}}\sup_{s\in(0,t]}(\sqrt{s}\|g_s\|_{H_1^1}).$$

If $C(\bar{W}_T) = C(T)$ actually does not depend on $\bar{W}_T$ we get for small enough t that

$$\sup_{s\in(0,t]}(\sqrt{s}\|g_s\|_{H_1^1}) \leq \frac{\|g_0\|_{L_1}}{1 - \sqrt{t}C(T)}$$

with a constant C implying (71). And in general we get a similar estimate a.s.

$\square$

18

Our basic objective is to study the sensitivity of the solution g_t with respect to initial data, that is

$$\xi_t(.;x)[g_0] = \frac{\delta g_t}{\delta g_0(x)} = \frac{d}{dh}|_{h=0} g_t[g_0 + h\delta_x]. \tag{75}$$

This can be done in general by analyzing the convergence of the successive approximations to the solutions

$$\xi_t^n(z;x)[g_0] = \frac{\delta \Phi_t^n[g_0](z)}{\delta g_0(x)},$$

which satisfies the recursion

$$\xi_t^n(z;x) = G(t,z,x) + \int_0^t ds \int \Big[G(t-s,z,y)\tilde{c}(s,y,[g_s])$$

$$- \frac{\partial}{\partial y}(G(t-s,z,y)\tilde{b}(s,y,[g_s])) \Big]\Big|_{g_s = \Phi_s^{n-1}[g_0]} \xi_s^{n-1}(y;x)\, dy$$

$$+ \int_0^t ds \iint \frac{\delta}{\delta g_s(w)} \Big[G(t-s,z,y)\tilde{c}(s,y,[g_s])$$

$$- \frac{\partial}{\partial y}(G(t-s,z,y)\tilde{b}(s,y,[g_s])) \Big]\Big|_{g_s = \Phi_s^{n-1}[g_0]} \xi_s^{n-1}(w;x)g_s(y)\, dy\, dw. \tag{76}$$

Under the assumptions of Theorem 4.1, say (i), we get the recursive estimates for ξ_n in the form

$$\|\xi_t^n(.;x)\|_{\mathcal{M}^{sign}(\mathbf{R})} \le C + C(T, \bar{W}_t, \sup_{t \in [0,T]} \|g_t\|_{L_1}) \int_0^t (t-s)^{-1/2} \|\xi_s^{n-1}\|_{\mathcal{M}^{sign}(\mathbf{R})} ds,$$

and by linearity the same estimates for the increments $\xi_t^{n+1}(.;x) - \xi_t^{n-1}(.;x)$ in terms of the increments $\xi_t^n(.;x) - \xi_t^{n-1}(.;x)$ implying the convergence of the sequence ξ^n and hence the existence of the derivative (75) almost surely.

To apply Theorem 4.1 to equation (40), we have to calculate the variational derivatives of the type $\delta F(\exp\{\Omega W_t\}g)/\delta g$ in terms of the derivatives of F. To this end, let us first find out, how the transformation $e^{t\Omega}$ acts on measures (rather than functions). For any functions $v \in L_1$, $\phi \in C(\mathbf{R})$ we have

$$(\phi, v) = \int \phi(x)e^{\Omega t}v(x)\, dx = \int G(t,x)\phi(x)v(Y(t,x))dx$$

$$= \int G(t, Y(-t,z))\phi(Y(-t,z))v(z)\frac{\partial Y(-t,z)}{\partial z}\, dz,$$

from which the extension to measures v is directly seen. Thus, for any measure g, we get

$$(\phi, \frac{\delta}{\delta g(x)}e^{t\Omega}g) = (\phi, e^{t\Omega}\delta_x) = G(t, Y(-t,x))\phi(Y(-t,x))\frac{\partial Y(-t,x)}{\partial x},$$

so that

$$\frac{\delta}{\delta g(x)}e^{t\Omega}g = G(t, Y(-t,x))\frac{\partial Y(-t,x)}{\partial x}\delta_{Y(-t,x)}.$$

Consequently,

$$\frac{\delta}{\delta g(x)} F(\exp\{\Omega W_t\}g) = \int \left.\frac{\delta F(\mu)}{\delta \mu(z)}\right|_{\mu=\exp\{\Omega W_t\}g} (e^{\Omega W_t}\delta_x)(z)\,dz$$

$$= \left.\frac{\delta F(\mu)}{\delta \mu(z)}\right|_{\substack{z=Y(-W_t,x)\\ \mu=\exp\{\Omega W_t\}g}} G(W_t, Y(-W_t,x))\frac{\partial Y(-W_t,x)}{\partial x}. \qquad (77)$$

Using this formula, equation (53) - (55), and the convergence of sequence (76), we obtain the following result as a consequence of Theorem 4.1.

Theorem 4.2. *Assume* (46) – (49) *hold and a $T > 0$ given. Then*

(i) For any $v_0 \in \mathcal{M}^{sign}(\mathbf{R})$ there exists a unique solution v_t of equation (40) on $[0,T]$ such that $v_t \in L_1(\mathbf{R})$ for all $t > 0$, positive whenever v_0 is positive, and

$$\|v_t\|_{L_1} \le C_1(T)\exp\{C_1(T)\bar{W}_T\}\|v_0\|_{\mathcal{M}(\mathbf{R})}, \quad \mathbf{E}\|v_t\|_{L_1} \le C_2(T)\|v_0\|_{\mathcal{M}(\mathbf{R})} \qquad (78)$$

with constants $C_{1,2}(T)$.

(ii) If $v_0 \in H_1^1$, then

$$\|v_t\|_{H_1^1} \le C_1(T)\exp\{C_1(T)\bar{W}_T\}\|v_0\|_{H_1^1}, \quad \mathbf{E}\|g_t\|_{H_1^1} \le C_2(T)\|g_0\|_{H_1^1}. \qquad (79)$$

(iii) For any $v_0 \in \mathcal{M}(\mathbf{R})$, $v_t \in H_1^1$ a.s. for all $t > 0$ and, if the bounds on the r.h.s. of (61) do not depend on $\bar{W}_T$, one has the estimate

$$\|v_t\|_{H_1^1} \le C_3(T)\frac{1}{\sqrt{t}}\|v_0\|_{\mathcal{M}(\mathbf{R})}. \qquad (80)$$

(iv) The variational derivative $\xi_t(.; x)[v_0] = \frac{\delta v_t}{\delta v_0(x)}$ of the solution v_t with respect to initial data exists a.s. as a measure of finite total variation.

5 Sensitivity for McKean-Vlasov SPDEs

We shall discuss in more detail the sensitivity of McKean-Vlasov SPDE (40) reducing our attention to a more specific case of L having the form of a dual second order operator, namely to the equation

$$dv = L'_{t,v}v\,dt - \nabla(A(x)v) \circ dW_t, \qquad (81)$$

where

$$L_{t,v}\phi = \frac{1}{2}\sigma^2(x)\frac{\partial^2\phi}{\partial x^2} + b(t,x,[v])\frac{\partial\phi}{\partial x}, \qquad (82)$$

and $L'_{t,v}$, its dual, defined as

$$L'_{t,v}u = \frac{1}{2}\frac{\partial^2}{\partial x^2}(\sigma^2(x)u(x)) - \frac{\partial}{\partial x}(b(t,x,[v])u(x)). \qquad (83)$$

Equation (81) is a particular case of (40), so that the theory of the previous section applies. Moreover, this equation naturally rewrites in the weak form as

$$d(\phi, v) = (L_{t,v}\phi, v)\,dt + (\Omega\phi, v) \circ dW_t, \qquad (84)$$

with $\Omega\phi = A(x)\nabla\phi$.

Making in (81) the change of function to $g = \exp\{-\Omega' W_t\}v$, where $\Omega' = -\nabla \circ A(x)$ is the dual to Ω, leads to the equation

$$\dot{g} = \exp\{-\Omega' W_t\}L'_{t,\exp\{\Omega' W_t\}g} \exp\{\Omega' W_t\}g \, dt, \qquad (85)$$

or in the weak form

$$\frac{d}{dt}(\phi, g) = (\tilde{L}_{t,\exp\{\Omega' W_t\}g}\phi, g) = (\exp\{\Omega W_t\}L_{t,\exp\{\Omega' W_t\}g}\exp\{-\Omega W_t\}\phi, g). \qquad (86)$$

Notice now that the operator $\Omega = A(x)\nabla$ coincides with (41) with vanishing B, and hence the corresponding transformation $e^{t\Omega}$ given by (50) has $G = 1$ and hence the estimates (57) does not contain $\bar{W}_T$. Moreover,

$$e^{t\Omega'}v(z) = v(Y(-t, z))\frac{\partial Y(-t, z)}{\partial z}, \qquad (87)$$

so that the estimate (57) for the operator Ω' also does not contain $\bar{W}_T$. Consequently Theorem 4.2 for equation (81) holds in its strongest form containing estimate (80). Moreover, general formulas (52)-(55) simplify essentially for $B = 0, c = 0$ allowing us to rewrite $\tilde{L}$ from (86) as

$$\tilde{L}_{t,\exp\{\Omega' W_t\}g}\phi = \frac{1}{2}\sigma^2(Y(W_t, x))\frac{\partial^2\phi}{\partial x^2} + \tilde{b}(t, x, [g])\frac{\partial\phi}{\partial x}, \qquad (88)$$

where $Y(t, x)$ solves the ODE $\dot{Y} = -A(Y)$ with the initial condition $Y(0, x) = x$ and

$$\tilde{b}(t, x, [g]) = \left(b(t, z, [\exp\{\Omega' W_t\}g])\frac{\partial Y}{\partial z}(-W_t, z) + \frac{1}{2}\sigma^2(z)\frac{\partial^2 Y}{\partial z^2}(-W_t, z)\right)\Bigg|_{z=Y(W_t, x)}. \qquad (89)$$

Furthermore, as operator (88) is the generator of a diffusion, its solution cannot increase the sup-norm, and hence the solution to equation (81) does not increase the L_1-norm (or, equivalently, the $\mathcal{M}(\mathbf{R})$-norm).

To study sensitivity of equation (81), we can now apply the results of [33] and [37] to equation (86). However, these results yield the existence of the derivatives with respect to initial data for almost all W_t, and we are interested here in the expectation of all bounds. Therefore, we sketch briefly the approach of [37] to see how the estimates for the expectation arise.

Let us differentiate (88) to get the equation for the derivatives

$$\xi_t(.; x)[g_0] = \frac{\delta g_t}{\delta g_0(x)} = \frac{d}{dh}|_{h=0}g_t[g_0 + h\delta_x].$$

the existence of these derivatives is already proved in Theorem 4.2.

Using (77) we get

$$(\phi, \dot{\xi}_t(.; x)[g_0]) = (\tilde{L}_{t,\exp\{\Omega' W_t\}g}\phi, \xi_t(.; x)[g_0])$$

$$+ \iint \left(\frac{\delta b(t, z, [\exp\{\Omega' W_t\}g])}{\delta g(r)}\frac{\partial Y(-W_t, z)}{\partial z}\right)\Bigg|_{z=Y(W_t, y)} \xi_t(r; x)[g_0]\frac{\partial\phi}{\partial y}g_t(y)dydr. \qquad (90)$$

Thus the evolution of ξ_t, considered as measures, is dual to the evolution on functions defined in the inverse time via the equation

$$\dot{\phi}_s = -\tilde{L}_{s,\exp\{\Omega' W_t\}g}\phi_s$$

$$-\int \left(\frac{\delta b(t, z, [\exp\{\Omega' W_t\}g])}{\delta g(.)} \frac{\partial Y(-W_t, z)}{\partial z} \right)\Bigg|_{z=Y(W_t,y)} \frac{\partial \phi_s}{\partial y} g_s(y)\, dy. \tag{91}$$

This equation defines the backward propagator $U^{s,t}$, $s \leq t$, on $C^1(\mathbf{R})$, such that $U^{s,t}\phi$ is the solution to equation (91) with the terminal condition ϕ at time t, and $U^{s,t} = (V^{t,s})'$, where $V^{t,s}$ is the forward propagator yielding the solution to equation (90). To see that $U^{s,t}$ is well defined as claimed, let us write (91) more explicitly. Namely, as follows from (87) and (77),

$$\frac{\delta}{\delta g(x)} F(\exp\{\Omega' W_t\}g) = \frac{\delta F(\mu)}{\delta \mu(z)}\Bigg|_{\mu=\exp\{\Omega' W_t\}g, z=Y(W_t,x)}. \tag{92}$$

Consequently, (91) rewrites as

$$\dot{\phi}_s(p) = -\tilde{L}_{s,\exp\{\Omega' W_t\}g}\phi_s(p)$$

$$-\int \frac{\delta b(t, z, \mu)}{\delta \mu(r)}\Bigg|_{\mu=\exp\{\Omega' W_t\}g_s}^{r=Y(W_t,p)} \frac{\partial Y(-W_t, z)}{\partial z}\Bigg|_{z=Y(W_t,p)} \frac{\partial \phi_s}{\partial y} g_s(y)\, dy. \tag{93}$$

From the form of $\tilde{L}$ it is seen that it generates a Feller semigroup in $C(\mathbf{R})$ with an invariant domain $C^1(\mathbf{R})$, and the second term of (93) is a bounded operator in $C^1(\mathbf{R})$, due to the assumption (49) on the norm of $\delta b(t, y, [v])/\delta v(.)$ in $C^1(\mathbf{R})$. Thus one can solve (93) by the standard perturbation theory showing that

$$\|U_{s,t}\phi\|_{C^1(\mathbf{R})} \leq C(T)\|\phi\|_{C^1(\mathbf{R})}$$

with $C(T)$ depending only on the $\|g_0\|_{\mathcal{M}(\mathbf{R})}$ and not on the noise implying that

$$\mathbf{E}\|U_{s,t}\phi\|_{C^1(\mathbf{R})} \leq C(T)\|\phi\|_{C^1(\mathbf{R})}.$$

Consequently the dual propagator $V^{t,s}$ defining the solution to ξ in equation (90) is a bounded propagator in the dual space $(C^1(\mathbf{R}))'$, both a.s. and on average. Similarly, assuming additional smoothness of coefficients we can claim that $U^{s,t}$ acts in $C^2(\mathbf{R})$. Finally, by moving the derivative of ϕ in the second term of (93) to g via the integration by parts and using (80) allows one to show that $U^{s,t}$ acts as a bounded semigroup in $C(\mathbf{R})$. Therefore, the estimates

$$\|U_{s,t}\phi\|_{C^k(\mathbf{R})} \leq C(T)\|\phi\|_{C^k(\mathbf{R})}, \quad \mathbf{E}\|U_{s,t}\phi\|_{C^k(\mathbf{R})} \leq C(T)\|\phi\|_{C^k(\mathbf{R})} \tag{94}$$

hold for $k = 0, 1, 2$ with constants $C(T)$ depending only on the $\|v_0\|_{\mathcal{M}(\mathbf{R})}$. This implies that the dual propagator $V^{t,s}$, solving equation (90), is a bounded propagator in the dual spaces $(C(\mathbf{R}))'$, $(C^1(\mathbf{R}))'$ and $(C^2(\mathbf{R}))'$ with bounds independent of the noise. Taking finally into account that

$$\xi_0(., x) = \delta_x \in \mathcal{M}(\mathbf{R}), \quad \frac{\partial}{\partial x}\xi_0(., x) = \delta'_x \in (C^1(\mathbf{R}))', \quad \frac{\partial^2}{\partial x^2}\xi_0(., x) \in (C^2(\mathbf{R}))',$$

leads us to the following result.

Theorem 5.1. *Let $T > 0$ and*

$$0 < \sigma_1 \le \sigma(x) \le \sigma_2, \quad 0 < A_1 \le A(x) \le A_2, \tag{95}$$

that $\sigma \in C^2(\mathbf{R}), A \in C^3(\mathbf{R})$ so that

$$\sup_x \max(|\sigma'(x)|, |\sigma''(x)|, |A'(x)|, |A''(x)|, |A'''(x)|) \le B_1. \tag{96}$$

Let

$$\|b(t, ., [v])\|_{C^2(\mathbf{R})} \le b_1, \tag{97}$$

$$\sup_{t,y} \sup_{\|v\|_{\mathcal{M}(\mathbf{R})} \le \lambda} \max \left(\left\| \frac{\delta b(t, y, [v])}{\delta v(.)} \right\|_{C^2(\mathbf{R})}, \left\| \frac{\delta}{\delta v(.)} \frac{\partial b(t, y, [v])}{\partial y} \right\|_{C(\mathbf{R})} \right) \le C(\lambda), \tag{98}$$

with some constants $\sigma_1, \sigma_2, A_1, A_2, B_1, b_1$ and a function $C(\lambda)$. Then the following holds:

(i) For any $v_0 \in \mathcal{M}^{sign}(\mathbf{R})$ there exists a unique solution v_t of equation (40) on $[0, T]$ such that $v_t \in L_1(\mathbf{R})$ for all $t > 0$, positive whenever v_0 is positive, and with the norm not exceeding $\|v_0\|_{\mathcal{M}(\mathbf{R})}$ for all realization of the noise W. Moreover, $v_t \in H_1^1$ for all $t > 0$ and the following estimates hold

$$\|v_t\|_{H_1^1} \le C(T)\|v_0\|_{H_1^1}, \quad \mathbf{E}\|v_t\|_{H_1^1} \le C(T)\|v_0\|_{H_1^1}, \tag{99}$$

$$\|v_t\|_{H_1^1} \le C(T)\frac{1}{\sqrt{t}}\|v_0\|_{\mathcal{M}(\mathbf{R})}, \quad \mathbf{E}\|v_t\|_{H_1^1} \le C(T)\frac{1}{\sqrt{t}}\|v_0\|_{\mathcal{M}(\mathbf{R})}. \tag{100}$$

(ii) The variational derivatives $\xi_t(.; x)[v_0] = \frac{\delta v_t}{\delta v_0(x)}$ of the solution v_t with respect to initial data are well defined as elements of $L_1(\mathbf{R})$ for any x and $t > 0$, and their first and second derivatives with respect to x are bounded elements of the dual spaces $(C^1(\mathbf{R}))'$ and $(C^2(\mathbf{R}))'$ respectively, so that

$$\|\xi_0(.; x)\|_{L_1} \le C(T), \quad \|\frac{\partial}{\partial x}\xi_0(.; x)\|_{(C^1(\mathbf{R}))'} \le C(T), \quad \|\frac{\partial^2}{\partial x^2}\xi_0(.; x)\|_{(C^2(\mathbf{R}))'} \le C(T) \tag{101}$$

with constants $C(T)$ depending only on the norm $\|v_0\|_{\mathcal{M}(\mathbf{R})}$ and independent of the noise.

We are also interested in the second derivatives of the solutions v_t with respect to initial data:

$$\eta_t(.; x_1, x_2) = \frac{d}{dh}\big|_{h=0}\xi_t(.; x_1)[v_0 + h\delta_x] = \frac{\partial^2}{\partial h_1 \partial h_2}\big|_{h_1=h_2=0}\mu_t[v_0 + h_1\delta_{x_1} + h_2\delta_{x_2}]. \tag{102}$$

For this η we get the following equation differentiating (90) with respect g_0;

$$(\phi, \dot{\eta}_t(.; x_1, x_2)) = (\tilde{L}_{t,\exp\{\Omega'W_t\}g}\phi, \eta_t(.; x_1, x_2))$$

$$+ \iint \left(\frac{\delta b(t, z, [\exp\{\Omega'W_t\}g])}{\delta g(r)} \frac{\partial Y(-W_t, z)}{\partial z} \right) \Big|_{z=Y(W_t,x)} \eta_t(r; x_1, x_2) \frac{\partial \phi}{\partial y} g_t(y) dy dr$$

$$+ \iint \left(\frac{\delta b(t, z, [\exp\{\Omega'W_t\}g])}{\delta g(r)} \frac{\partial Y(-W_t, z)}{\partial z} \right) \Big|_{z=Y(W_t,x)} \frac{\partial \phi}{\partial y}\tilde{\xi}_t(r, y; x_1, x_2) dy dr \tag{103}$$

$$+ \iint \left(\frac{\delta^2 b(t, z, [\exp\{\Omega' W_t\} g])}{\delta g(r_1) \delta g(r_2)} \frac{\partial Y(-W_t, z)}{\partial z} \right) \bigg|_{z = Y(W_t, x)} \frac{\partial \phi}{\partial y} \xi_t(r_2; x_2) \xi_t(r_1; x_1) g_t(y) \, dy dr_1 dr_2,$$

where $\tilde{\xi}_t(r, y; x_1, x_2) = [\xi_t(r; x_2) \xi_t(y; x_1) + \xi_t(y; x_2) \xi_t(r; x_1)]$.

The well-posedness of this equations and then the existence of the derivative (102) follows as above. However, we need also the existence and bounds for the derivatives of η with respect to x_1, x_2.

Theorem 5.2. *Under assumption of Theorem 5.1 let $b \in C^{2,1\times1}(\mathcal{M}_1^{sign}(\mathbf{R}))$ as a function of v with all bounds uniform in other variables. Then for any $v_0 \in \mathcal{M}_1(\mathbf{R})$ the derivative (102) is well defined for all $t > 0, x$ and the following bounds hold*

$$\| \frac{\partial^\alpha}{\partial x_1^\alpha} \frac{\partial^\beta}{\partial x_2^\beta} \eta_t(.; x_1, x_2) \|_{[C^2(\mathbf{R})]'} \leq C(T) \tag{104}$$

with $\alpha, \beta \leq 1$ and with some (random) constants C depending on the time horizon T, but not on the noise W.

Proof. In the light of the properties of ξ from Theorem 5.1 it is straightforward to see, differentiating equation (103) with respect to x_1 and x_2 that the assumptions made on b is precisely the one needed to make all terms not containing η uniformly bounded (due to the product structure of ξ entering the equation for η), so that they can be written by the usual perturbation arguments. $\square$

6 On the domain of the Markov semigroups generated by the McKean-Vlasov SPDEs

Since equation (81) its solutions defines a Markov process, in fact a measure-valued diffusion, the corresponding Markov propagator being given on the continuous functionals of measures in the usual way:

$$U^{s,t} F(v) = \mathbf{E} F(v_t(v, [W])), \tag{105}$$

where v_t is the solution to (81) for $t > s$ with given $v = v_s$ at time s.

We use the same letter U that was used for the propagators discussed in the proof if Theorem 5.1 which should not cause any confusion, as U is used in the sense of (105) everywhere, except in the intermediate discussion leading to Theorem 5.1.

The main conclusion we need from the sensitivity analysis developed above is the invariance of the set of smooth functionals under this propagator, that is the following fact:

Theorem 6.1. *Under assumption of Theorem 5.2 the spaces of functionals $C^{1,2}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))$ and its intersection with $C^{2,1\times1}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))$ are invariant under the action of the operators (105), so that*

$$\|U^{s,t} F\|_{C^{1,2}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))} \leq C(T) \|F\|_{C^{1,2}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))}, \tag{106}$$

$$\|U^{s,t} F\|_{C^{2,1\times1}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))} \leq C(T) \left(\|F\|_{C^{2,1\times1}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))} + \|F\|_{C^{1,2}(\mathcal{M}_\lambda^{sign}(\mathbf{R}))} \right) \tag{107}$$

with a constant $C(T)$.

Proof. It follows from Theorem 5.1 and the formula

$$\frac{(\delta U^{s,t}(F))(v)}{\delta v(x)} = \mathbf{E}\frac{\delta F(v_t)}{\delta v(x)} = \mathbf{E}\int_{\mathbf{R}}\frac{\delta F(v_t)}{\delta v_t(z)}\xi_t(z;x)[v]\,dz,$$

that

$$\|U^{s,t}F\|_{C^{1,2}(\mathcal{M}^{sign}_\lambda(\mathbf{R}))} \leq \mathbf{E}\left\|\int_{\mathbf{R}}\frac{\delta F(v_t)}{\delta v_t(z)}\xi_t(z;.)[v]\,dz\right\|_{C^2(\mathbf{R})}$$

$$\leq \mathbf{E}\left\|\frac{\delta F(v_t)}{\delta v_t(.)}\right\|_{C^2(\mathbf{R})}\left\|\frac{\partial^2}{\partial x^2}\xi_t(.;x)[v]\right\|_{(C^2(\mathbf{R}))'} \leq C(T).$$

It follows from Theorem 5.2 and the formula

$$\frac{(\delta^2 U^{s,t}(F))(v)}{\delta v(x)\delta v(y)} = \mathbf{E}\frac{\delta^2 F(v_t)}{\delta v(x)\delta v(y)}$$

$$= \mathbf{E}\int_{\mathbf{R}}\frac{\delta F(v_t)}{\delta v_t(z)}\eta_t(z;x,y)[v_0]dz + \mathbf{E}\int_{\mathbf{R}^2}\frac{\delta^2 F(v_t)}{\delta v_t(z)\delta v_t(w)}\xi_t(z;x)[v_0]\xi_t(w;y)[v_0]\,dzdw,$$

that

$$\|U^{s,t}F\|_{C^{2,1\times 1}(\mathcal{M}^{sign}_\lambda(\mathbf{R}))} \leq \mathbf{E}\left\|\frac{\delta^2 F(v_t)}{\delta v(.)\delta v(.)}\right\|_{C^{1\times 1}(\mathbf{R}^2)}$$

$$\leq \left\|\frac{\partial^\alpha}{\partial x_1^\alpha}\frac{\partial^\beta}{\partial x_2^\beta}\eta_t(.;x_1,x_2)\right\|_{[C^2(\mathbf{R})]'}\|F\|_{C^{1,2}(\mathcal{M}^{sign}_\lambda(\mathbf{R}))}$$

$$+ \left\|\frac{\partial^\alpha}{\partial x^\alpha}\xi_t(.;x)[v]\right\|_{(C^1(\mathbf{R}))'}\left\|\frac{\partial^\beta}{\partial y^\beta}\xi_t(.;y)[v]\right\|_{(C^1(\mathbf{R}))'}\left\|\frac{\delta^2 F(v_t)}{\delta v_t(.)\delta v_t(.)}\right\|_{C^{2,1\times 1}(\mathcal{M}^{sign}_\lambda(\mathbf{R}))},$$

leading to (107). $\qquad\square$

Acknowledgement

References

[1] Crisan, D., Kurtz, Th., Lee, Y. (2014). Conditional distributions, exchangeable particle systems, and stochastic partial differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* 50(4):946–974.

[2] Kolokoltsov, V.N. (2013). Nonlinear diffusions and stable-like processes with coefficients depending on the median or VaR. *Appl. Math. Optim.* 68(1):85–98.

[3] Kunita, H. (1997). *Stochastic flows and stochastic differential equations.* Cambridge Stud. Adv. Math. Cambridge, UK: Cambridge University Press.

[4] Kurtz, Th., Xiong, J. (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic. Process. Appl.* 83:103–126.

[5] Cardaliaguet, P., Delarue, F., Lasry, J.-M., Lions, P.-L. (2015) The master equation and the convergence problem in mean field games. Available at: https://arxiv.org/abs/1509.02505.

[6] Carmona, R., Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I. Mean Field FBSDEs, Control, and Games.* Probab. Theory Stoch. Model. 83.

[7] Carmona, R., Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications II. Mean Field Games with Common Noise and Master Equations.* Probab. Theory Stoch. Model. 84.

[8] Huang, M., Malhamé, R., Caines, P. (2006). Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* 6:221–252.

[9] Nourian, M., Caines, P. (2013). ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM J. Control Optim.* 51(4):3302–3331.

[10] Kolokoltsov, V., Yang, W. (2013). Existence of solutions to path-dependent kinetic equations and related forward-backward systems. *Open J. Optim.* 2(2):39–44. DOI: 10.4236/ojop.2013.22006.

[11] Kolokoltsov, V., Yang, W. (2013). Sensitivity analysis for HJB equations with an application to a coupled backward-forward system. Available at: https://arxiv.org/abs/1303.6234.

[12] Carmona, R., Delarue F., Lacker, D. (2016). Mean field games with common noise arXiv:1407.6181 or arXiv:1407.6181v2. *Ann. Probab.* 44(6):3740–3803.

[13] Lasry, J.-M., Lions, P.-L. (2006) Jeux à champ moyen. I. Le cas stationnaire (French). *Comptes Rendus Math?matique.* 343(9):619–625.

[14] Huang, M., Caines, P., Malhamé, R. (2007). Large-Population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized ϵ-Nash Equilibria. *IEEE Trans. Automat. Control.* 52(9):1560–1571.

[15] Huang, M. (2010) Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.* 48:3318–3353.

[16] Bardi, M., Caines, P., Capuzzo Dolcetta, I. (2013). Preface: DGAA special issue on mean field games. *Dyn. Games Appl.* 3(4):443–445.

[17] Bensoussan, A., Frehse, J., Yam, Ph. (2013). *Mean field games and mean field type control theory.* Springer Briefs in Mathematics. NY: Springer.

[18] Guéant, O., Lasry, J.-M., Lions, P.-L. (2011) Mean Field Games and Applications. In: *Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Math. 2003* Springer, Berlin, Heidelberg, pp. 205–266.

[19] Gomes, D.A., Saude, J. (2014). Mean field games models–a brief survey. *Dyn. Games Appl.* 4(2):110–154.

[20] Caines, P.E. (2015) Mean Field Games. In: Baillieul, J., Samad, T., eds. *Encyclopedia of Systems and Control.* London, UK: Springer, pp. 706–711.

[21] Cardaliaguet, P., Lasry, J.-M., Lions, P.-L., Porretta, A. (2013). Long time average of mean field games with a nonlocal coupling. *SIAM J. Control Optim.* 51(5):3558–3591.

[22] Carmona, R., Lacker, D. (2015) A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.* 25(3):1189–1231.

[23] Carmona, R., Delarue, F. (2013). Probabilistic analysis of mean-field games. *SIAM J. Control Optim.* 514:2705–2734.

[24] Gomes, D.A., Patrizi, S., Voskanyan, V. (2014). On the existence of classical solutions for stationary extended mean field games. *Nonlinear Anal.* 99:49–79.

[25] Achdou, Y., Camilli, F., Capuzzo-Dolcetta, I. (2013). Mean field games: convergence of a finite difference method. *SIAM J. Numer. Anal.* 51(5):2585–2612.

[26] Tembine, H., Zhu, Q., Basar, T. (2014). Risk-sensitive mean-field games. *IEEE Trans. Automat. Control.* 59(4):835–850.

[27] Gomes, D., Velho, R.M., Wolfram, M.-T. (2014). Socio-economic applications of finite state mean field games. *Philos. Trans. Roy. Soc. A.* 372(2028):20130405.

[28] Carmona, R., Fouque, J.-P., Sun, L.-H. (2015). Mean field games and systemic risk. *Commun. Math. Sci.* 13(4):911–933.

[29] Ahuja, S. (2016) Wellposedness of mean field games with common noise under a weak monotonicity condition. *SIAM J. Control Optim.* 54(1):30–48.

[30] Lacker, D. (2016). A general characterization of the mean field limit for stochastic differential games. *Probab. Theory Related Fields* 165(3-4):581–643.

[31] Bensoussan, A., Frehse, J., Yam, S.C.P. (2014). The master equation in mean field theory. *Journal de Mathematiques Pures et Appliquees.* 103(6):1441–1474.

[32] Dawson, D., Vaillancourt, J. (1995). Stochastic McKean-Vlasov equations. *NoDEA* 2:199–229.

[33] Kolokoltsov, V.N. (2010). *Nonlinear Markov processes and kinetic equations.* Cambridge Tracks in Mathematics, 182, Cambridge, UK: Cambridge University Press.

[34] Hu, Y., Nualart, D., Song, J. (2013). A nonlinear stochastic heat equation: Hölder continuity and smoothness of the density of the solution. *Stochastic Process. Appl.* 123(3):1083–1103.

[35] Conus, D., Joseph, M., Khoshnevisan, D. (2013). On the chaotic character of the stochastic heat equation, before the onset of intermittency. *Ann. Probab.* 41(3B):2225–2260.

[36] Kolokoltsov, V.N., Tyukov, A.E. (2003). Small time and semiclassical asymptotics for stochastic heat equation driven by Lévy noise. *Stoch. Stoch. Rep.* 75(1-2):1–38.

[37] Kolokoltsov, V.N., Troeva, M., Yang, W. (2014). On the rate of convergence for the mean-field approximation of controlled diffusions with large number of players. *Dyn. Games Appl.* 4(2):208–230.

[38] Ito, S. (1992). Diffusion equatons. In: *Transactions of Mathematical Monographs*, Vol. 114, AMS, Providence, RI.

[39] Porper, F.O., Eidelman, S.D. (1984). Two-sided estimates of the fundamental solutions of second-order parabolic equations and some applications of them. (Russian) *Uspekhi Mat. Nauk.* 39(3):107–156.