Strong well-posedness of McKean-Vlasov stochastic differential equation with Hölder drift

Paul-Éric Chaudru de Raynal

STRONG WELL POSEDNESS OF MCKEAN-VLASOV STOCHASTIC DIFFERENTIAL EQUATIONS WITH HÖLDER DRIFT

P.E. CHAUDRU DE RAYNAL

Abstract. Here, we prove strong well-posedness for stochastic systems of McKean-Vlasov type with Hölder drift, even in the measure argument, and uniformly non-degenerate Lipschitz diffusion matrix. The Hölder regularity of the drift with respect to the law argument being for the Wasserstein distance.

Our proof is based on Zvonkin's transformation (1974) and so on the regularization properties of the associated PDE, which is stated on the space $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, where $T>0$, d denotes the dimension equation and $\mathcal{P}_2(\mathbb{R}^d)$ is the space of probability measures on $\mathbb{R}^d$ with finite second order moment. Especially, a smoothing effect in the measure direction is exhibited. Our approach is based on a parametrix expansion of the transition density of the McKean-Vlasov process.

1 Introduction

Let $\mathcal{M}_d(\mathbb{R})$ be the set of $d\times d$ matrices with real coefficients and $\mathcal{P}_2(\mathbb{R}^d)$ be the space of probability measures ν on $\mathbb{R}^d$ such that $\int x^2\mathrm{d}\nu(x)<+\infty$. For any random variable Z, let us denote by $[Z]$ its law.

For a positive number T, for given measurable functions $B,\Sigma:[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d\times\mathcal{M}_d(\mathbb{R})$ and for $(B_t,t\geq 0)$ a standard d-dimensional Brownian motion defined on a filtered probability space $(\Omega,\mathcal{F},\mathbb{P},(\mathcal{F}_t)_{t\geq 0})$, we consider, for $t<s$ in $[0,T]^2$ and μ in $\mathcal{P}_2(\mathbb{R}^d)$, the non-linear (in a McKean-Vlasov sense) system

$$X_s^{t,\mu}=X_t+\int_t^s B(r,X_r^{t,\mu},[X_r^{t,\mu}])\mathrm{d}r+\int_t^s\Sigma(r,X_r^{t,\mu},[X_r^{t,\mu}])\mathrm{d}B_r,\quad X_t\sim\mu. \tag{1.1}$$

This sort of equation arises as the limit of system of interacting players. This happens as follows. Suppose that we are given a large number of players with symmetric dynamic and whose positions depend on the positions of the other players in a mean field way. Then, when the number of players tends to infinity, there is a propagation of chaos phenomenon so that the limit dynamic of each player does not depend on the positions of the others anymore, but only on their statistical distributions. This obviously comes from the law of large numbers. The resulting system is then of the form of (1.5) and is called non-linear, since the dynamic of the player depends on its own law. We refer to the notes of Sznitman's lecture at Saint-Flour [Szn91] for an overview on the topic.

As done in [Szn91], the proof of strong existence and uniqueness (which means that the solutions are adapted to the filtration generated by the Brownian motion and are almost surely indistinguishable) for this equation relies on classical fixed point argument and so, on the Lipschitz property of the

Key words and phrases. McKean-Vlasov processes; smoothing effect; non-linear PDE; regularization by noise.

coefficients of the equation, the Lispchitz regularity being understood with respect to (w.r.t.) the Wasserstein metric in the case of the measure argument.

In this work, we aim at proving the strong well-posedness of such a system when the drift B is a measurable and only a bounded and Hölder continuous in both the space and measure arguments and when the diffusion matrix Σ is non-degenerate and Lipschitz in the space and measure arguments. The Hölder assumption w.r.t. the measure argument has to be understood as follows : there exists a positive α_1 such that the drift function B is supposed to be bounded and measurable and α_1-Hölder w.r.t. the measure argument in the following sense:

$$\forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}^d, \forall \nu, \mu \in \mathcal{P}_2(\mathbb{R}^d) : \ |B(t,x,\nu) - B(t,x,\mu)| \leq \inf_{Z\sim\nu,\ Z'\sim\mu} \mathbb{E}|Z - Z'|^{\alpha_1}, \tag{1.2}$$

which can be seen as an α_1-Hölder regularity w.r.t. the Wasserstein distance.

Our proof is based on the regularization by noise phenomenon (see [Fla11] for a survey).This effect comes from the random perturbation of the equation, allowing stochastic differential system to be well-posed, in a strong or a weak sense, under a larger set of assumptions than ordinary differential system. It thus deeply relies on the noise propagation through the system which is the reason why we suppose the matrix $\Sigma\Sigma^*$ to be uniformly non-degenerate.

To the best of our knowledge, this result is new and then consists, on the one hand, to show that this phenomenon still holds in the McKean-Vlasov setting (which seems quite reasonable) and, on the other hand, even occurs in the measure space. This last result is quite unexpectable at first sight, since the noise does not act in that direction. In order to recover the regularization by noise phenomenon w.r.t. the measure argument without adding noise in the measure space, we have to require some structural assumptions on the dependence of the drift and diffusion coefficients w.r.t. the measure argument: the dependence upon the measure is supposed to be of polynomial type. Here, this assumption translates into the following condition: there exists two functions $b, \sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d \times \mathcal{M}_d(\mathbb{R})$, and two functions $\varphi_i : \mathbb{R}^d \to \mathbb{R}$, i in $\{1,2\}$ such that

$$\forall (t,x,\nu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d), \quad B(t,x,\nu) = b(t,x,\langle \varphi_1, \nu\rangle) \text{ and } \Sigma(t,x,\nu) = \sigma(t,x,\langle \varphi_2, \nu\rangle), \tag{1.3}$$

where, for any measurable function φ we wrote $\int \varphi \mathrm{d}\nu$ with its dual notation: $\langle \varphi, \nu\rangle$. Hence, the assumption (1.2) on the drift translates as follow: the function b is assumed to be bounded and measurable, Hölder continuous w.r.t. the second argument and Lipschitz w.r.t. the third argument but the map φ_1 is supposed to be only Hölder-continuous for some $\alpha_1 > 0$.

Remark 1. *Our results also hold for others polynomial dependence. As an example, instead of* (1.3) *we could have suppose that*

$$\forall (t,x,\nu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d), \quad B(t,x,\nu) = \langle b(t,x,\cdot), \nu\rangle \text{ and } \Sigma(t,x,\nu) = \langle \sigma(t,x,\cdot), \nu\rangle. \tag{1.4}$$

In that case, the function b has to be measurable and bounded and Hölder continuous w.r.t. the first and second space variable and the non-degeneracy assumption on the diffusion matrix has to be understood for the matrix $[\langle \sigma(t,x,\cdot), \nu\rangle]\,[\langle \sigma(t,x,\cdot), \nu\rangle]^*$ uniformly in (t,x) in $\mathbb{R}^+ \times \mathbb{R}^d$ and ν in the space of probability measures on $\mathbb{R}^d$.

Assuming the structural assumption (1.3), we can then rewrite our McKean-Vlasov system (1.1) as follow:

$$X_s^{t,\mu} = X_t + \int_t^s b(r, X_r^{t,\mu}, \langle \varphi_1, [X_r^{t,\mu}]\rangle)\mathrm{d}r + \int_t^s \sigma(r, X_r^{t,\mu}, \langle \varphi_2, [X_r^{t,\mu}]\rangle)\mathrm{d}B_r, \quad X_t \sim \mu. \tag{1.5}$$

To prove the strong well-posedness outside the Lipschitz framework, we adapt the Zvonkin transformation [Zvo74] to our non-linear case. This approach relies on smoothing properties of a well-chosen PDE associated to the system (1.5). Here, the non-linearity (in a Mckean-Vlasov sense) leads to a particular class of PDE that can be seen as the "linear"[1] version of the so called Master Equation coming from Mean Field Games theory introduced independently by Lasry and Lions [LL06b, LL06a, LL07] and by Huang, Caines and Malhameé [HMC06]. This PDE has been recently studied from a probabilistic point of view in the independent works of Buckdahn, Li, Peng and Rainer in [BLPR14] and of Crisan, Chassagneux and Delarue in [CCD14]. Its main particularity comes from the fact that it is stated on the space $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ so that it involves derivatives in the measure direction.

The smoothing properties of the associated PDE is then the crucial part of the proof. We investigate it under a larger set of assumptions: we let the diffusion matrix be only a Hölder-continuous function of the space and measure argument. The investigations are done by using a Feynman-Kack representation of the solution of the PDE and then a parametrix expansion (see [MS67]) of the transition density of the solution of (1.5). This brings us to investigate for all $t < s$ in $[0,T]^2$ the regularity of $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto \langle \phi, [X_s^{t,\mu}] \rangle \in \mathbb{R}$ for some Hölder continuous function $\phi : \mathbb{R}^d \to \mathbb{R}$ and where X is a solution of (1.5). Especially, as it has been announced above, a smoothing effect w.r.t. measure argument (*i.e.* on the initial data "μ") is exhibited.

Finally, let us emphasize that the parametrix approach relies on perturbation approach. This explains why, in comparison with the results obtained in the linear case (see [Zvo74, Ver80]), the function φ_1 appearing in the drift coefficients of (1.5) is assumed to be Hölder continuous and not only a bounded function.

Organization of this paper.

Our paper is organized as follows: we present below our main assumptions and result. Then, we give in Section 2 the mathematical background and our strategy of proof. Especially, we state in this section the PDE associated to (1.5). Since the PDE is stated on the Cartesian space $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, we give some notions on differentiation of functions along a measure. Then, we establish the smoothing properties of the PDE associated to (1.5), which is a key result in the proof of our main Theorem.

Next, we investigate the smoothing properties of the PDE. Such an investigation is done under regularized framework. It is based on a parametrix expansion of the transition density of (1.5) and is presented as follows. In Section 4 we give estimates on the transition density of (1.5) and on the mapping $v : \mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto \langle \phi, [X_s^{t,\mu}] \rangle \in \mathbb{R}$ for some Hölder continuous function $\phi : \mathbb{R}^d \to \mathbb{R}$. This permits to estimate the solution of the PDE. Then, estimates on v are proven in Section 5 and estimates on the transition density of (1.5) are proven in Section 6. As said before, this last follows from a parametrix representation of the transition density of (1.5) and its estimation. Auxiliary results are given in Appendixes A and B.

[1] *i.e.* without non-linear source term

Notations, assumptions and main result

Notations. For any function $f : E \times F \times G \to \mathbb{R}^N$, we denote by ∂_1 (resp. ∂_2 and ∂_3) the differentiation w.r.t. the first (resp. second and third) variable. When we add a subscript in the operator $\partial_\cdot$, it stands for the variable on which the differentiation operator acts. We recall that the law of a random variable X is denoted by $[X]$. The superscript "$*$" stands for the transpose, the canonical Euclidean inner product on $\mathbb{R}^d$ is denoted by "$\cdot$". We denote by $\mathcal{M}_d(\mathbb{R})$ the set of $d \times d$ matrices with real coefficients and the trace of a matrix M in $\mathcal{M}_d(\mathbb{R})$ is denoted by $\mathrm{Tr}(M) = \sum_{j=1}^d M_{jj}$. We let $C, C', c, c', \tilde{C}, \tilde{C}', \ldots$ be some positive constants depending only on known parameters in **(HE)** given below, that may change from line to line and from an equation to another and we add a subscript T in the constant if it depends also on the length of the interval.

Assumptions (HE). We say that assumptions **(HE)** hold if the following assumptions are satisfied:

(HE1) regularity of the drift: there exists a positive constant C_b such that $||b||_\infty < C_b$. Moreover the mapping b is supposed to be Hölder continuous in space, uniformly in time and w.r.t. the law argument, for some Hölder exponent $\eta > 0$. For all (t, x) in $[0, T] \times \mathbb{R}^d$, the mapping $b(t, x, \cdot) : \mathbb{R} \ni w \mapsto b(t, x, w)$ is differentiable and $||\partial_3 b||_\infty < C'_b$. Finally, the mapping $\varphi_1 : \mathbb{R}^d \ni x \mapsto \varphi_1(x)$ is supposed to be α_1-Hölder for some $0 < \alpha_1 \leq 1$.

(HE2) regularity of the diffusion matrix: there exists a positive constant C_σ such that for all t in $[0, T]$,

$$\forall x, x' \in \mathbb{R}^d,\ w, w' \in \mathbb{R},\ |\sigma(t, x, w) - \sigma(t, x', w')| \leq C_\sigma \left(|x - x'| + |w - w'|\right).$$

Moreover for all (t, x) in $[0, T] \times \mathbb{R}^d$, the mapping $\sigma(t, x, \cdot) : \mathbb{R} \ni w \mapsto \sigma(t, x, w)$ is differentiable, $||\partial_3 \sigma||_\infty < C'_\sigma$ and there exists a positive constant C''_σ such that for all t in $[0, T]$ and w in $\mathbb{R}$,

$$\forall x, x' \in \mathbb{R}^d,\ |\partial_3 \sigma(t, x, w) - \partial_3 \sigma(t, x', w)| \leq C''_\sigma |x - x'|^{\gamma'_a}.$$

Finally, the mapping $\varphi_2 : \mathbb{R}^d \ni x \mapsto \varphi_2(x)$ is supposed to be Lipschitz.

(HE3) uniform ellipticity of $\sigma\sigma^*$: the function $\sigma\sigma^*$ satisfies the uniform ellipticity hypothesis:

$$\exists \Lambda > 1,\ \forall \zeta \in \mathbb{R}^d, \quad \Lambda^{-1}|\zeta|^2 \leq [\sigma\sigma^*(t, x, w)\zeta] \cdot \zeta \leq \Lambda |\zeta|^2,$$

for all $(t, x, w) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$.

Remark 2. *We emphasize that under **(HE)** the drift function b is not Lipschitz w.r.t. the law argument since the function φ_1 is only Hölder continuous.*

Main result. We can now state our main result:

Theorem 1.1. *Under assumptions **(HE)**, the system* (1.5) *admits a unique strong solution.*

2 Mathematical background and strategy of proof

2.1 The Zvonkin transformation

For usual differential equations, it could be a very hard task to show the well posedness outside the Lipschitz framework, at least in the classical sense (see [DL89] for some work in that direction). Nevertheless, when the differential system is perturbed by noise there is a phenomenon, called *regularization by noise*, that allows to recover the well posedness. When the SDE is linear (in a Mckean-Vlasov sense), this has been studied first by Zvonkin [Zvo74] and then generalized by several authors *e.g.* [Ver80, KR05, Zha11] and [Fla11] for a survey. All these results rely on smoothing properties of an associated PDE and so on smoothing properties of elliptic and linear partial second order differential operator. Let us briefly explain how.

The strategy to recover the Lipschitz property consists in exhibiting a Zvonkin-like transformation of the equation. Let us forget for the moment the dependence of the solution of (1.5) w.r.t. its own law in order to illustrate the main argument. If we denote by $\mathcal{A}$ the generator of (1.5), the idea is to obtain *a priori* estimates on the solution of the PDE

$$\partial_t \mathbf{u} + \mathcal{A}\mathbf{u} = b, \text{ on } [0,T) \times \mathbb{R}^d, \quad \mathbf{u}_T = 0_{\mathbb{R}^d}, \tag{2.1}$$

when b and σ are smooth functions, but depending only on regularity of b, σ assumed in **(HE)**.

This allows to consider a sequence $(\mathbf{u}^n)_{n\geq 0}$ of classical solutions of the PDE (2.1) along a sequence of mollified coefficients $((\sigma\sigma^*)_n, b_n)_{n\geq 0}$. Then, by applying Itô's formula on $X_t - \mathbf{u}^n(t, X_t)$ we can remove the drift of the equation and recover an SDE whose coefficients have Lipschitz constants uniformly on the regularization procedure, so that, when letting the regularization procedure tend to infinity, the estimates pass through the limit.

When these constants can be chosen as small as T is small (which follows from the boundary condition in (2.1)), we then recover existence and uniqueness on small time intervals. If in addition the constants do not degenerate with the time, we can iterate the procedure and then recover existence and uniqueness on $\mathbb{R}^+$.

The smoothing properties of the PDE (2.1), are, in fact, the crucial points. It is well known that such smoothing properties are related to the noise propagation in the associated SDE through all the directions of the space. Hence, two issues arise from the non linear framework studied here: how the operator $\mathcal{A}$ looks like in our Mckean-Vlasov case, and how to regularize in the measure direction since the noise does not act in that direction.

2.2 PDE on space of probability measure

Roughly speaking, we have to find a PDE that reflects the Markov structure of the underlying process. Here, the Markov property has to be understood on the space $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, so that it seems natural to consider a PDE on this space. This sort of PDE has been recently studied independently by Buckdahn, Li, Peng and Rainer in [BLPR14] and Crisan, Chassagneux and Delarue in [CCD14], it is called the Master Equation and it appears naturally when considering Mean-Field Games. What follows is essentially inspired by the second work [CCD14], from which we adopted some

of the notations.

Before giving this PDE, stated on the space $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, let us give some notions of differentiation of functions along a probability measure. The one used here has been introduced by Lions during its lecture at the *Collège de France* and can be found in Cardaliaguet's note [Car10]. The strategy of Lions consists in lifting the function $V:\mathcal{P}_2(\mathbb{R}^d)\ni\nu\mapsto V(\nu)\in\mathbb{R}$ to a function $\mathcal{V}:\mathbb{L}_2(\Omega,\mathcal{F},\mathbb{P})\ni Z\mapsto\mathcal{V}(Z)\in\mathbb{R}$, Z being a random variable of law ν. We can then take advantage of the Hilbert structure of the $\mathbb{L}_2$ space and define, in the Frechet sense, the mapping $D\mathcal{V}$. Thanks to Riezs' representation Theorem, we can identify $D\mathcal{V}(Z)$ as $DV(\nu)(Z)$. Thus, we call the derivative of V w.r.t. the law, and we denote by $\partial_\nu V(\nu)$, the mapping in $\mathbb{L}_2(\mathbb{R}^d,\nu;\mathbb{R}^d)$:

$$\partial_\nu V(\nu):\mathbb{R}^d\ni z\mapsto\partial_\nu V(\nu)(z)\in\mathbb{R}^d.$$

Let us emphasize that, in our case, the law interaction appears as the action of the law on some function φ_i : $V(\nu)=\int\varphi_i(x)\mathrm{d}\nu(x)$. Using the lifting argument described above we get that for any random variable X and H in $\mathbb{L}_2(\Omega)$:

$$V([X+\epsilon H])=\mathbb{E}[\varphi(X+\epsilon H)]=\mathbb{E}[\varphi(X)]+\epsilon\mathbb{E}[\varphi'(X)\cdot H]+o(\epsilon),$$

so that

$$\partial_\nu\langle\varphi_i,\nu\rangle:\mathbb{R}^d\ni z\mapsto\varphi_i'(z)\in\mathbb{R}^d.$$

Finally, let us just notice that this definition justifies the choice of the space $\mathcal{P}_2(\mathbb{R}^d)$ for the initial data in (1.5).

We can now state the PDE of interest:

$$\begin{cases}(\partial_t+\mathcal{A})\mathbf{u}(t,x,\mu)=b(t,x,\langle\varphi_1,\mu\rangle),\ \text{on}\ [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d).\\ \mathbf{u}(T,x,\mu)=0_{\mathbb{R}^d},\end{cases}\tag{2.2}$$

where, when setting $a:=\sigma\sigma^*$ the operator $\mathcal{A}$ is given by: for any smooth enough function $\psi:\mathbb{R}^+\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d$

$$\begin{aligned}\mathcal{A}\psi(t,x,\mu)\ &=\ \frac{1}{2}\mathrm{Tr}\left[a(t,x,\langle\mu,\varphi_2\rangle)\partial_x^2\psi(t,x,\mu)\right]+b(t,x,\langle\varphi_1,\mu\rangle)\partial_x\psi(t,x,\mu)\\ &\quad+\int b(t,z,\langle\varphi_1,\mu\rangle)\partial_\mu\psi(t,x,\mu)(z)\mathrm{d}\mu(z)+\frac{1}{2}\int\mathrm{Tr}\left[a(t,z,\langle\varphi_2,\mu\rangle)\partial_z(\partial_\mu\psi(t,x,\mu)(z))\right]\mathrm{d}\mu(z).\end{aligned}$$

When the coefficients are smooth, it follows from [CCD14] that such a PDE admits a classical solution $\mathbf{u}$. We refer to the aforementioned paper for more explanations on the meaning of classical solution and especially on the question about the regularity of $\partial_\mu\mathbf{u}$ as an element of $\mathbb{L}_2(\mathbb{R}^d,\mu;\mathbb{R}^d)$ w.r.t. the variable z and μ.

2.3 Smoothing properties of the PDE

It thus remains to show the smoothing properties of the PDE. This is done under the following assumptions.

Assumptions (**H**$\mathcal{E}$). We say that assumptions (**H**$\mathcal{E}$) hold if Assumptions **(HE)** are satisfied with assumption (**HE2**) replaced by

(H$\mathcal{E}$2) regularity of the diffusion matrix: there exists a positive constant C_σ such that for all t in $[0,T]$, for all w in $\mathbb{R}$

$$\forall x, x' \in \mathbb{R}^d,\ |\sigma(t,x,w) - \sigma(t,x',w)| \leq C_\sigma |x - x'|^{\gamma_a},$$

for some $0 < \gamma_a \leq 1$. Moreover for all (t,x) in $[0,T] \times \mathbb{R}^d$, the mapping $\sigma(t,x,\cdot) : \mathbb{R} \ni w \mapsto \sigma(t,x,w)$ is differentiable, $||\partial_3 \sigma||_\infty < C'_\sigma$ and there exists a positive constant C''_σ such that for all t in $[0,T]$ and w in $\mathbb{R}$,

$$\forall x, x' \in \mathbb{R}^d,\ |\partial_3\sigma(t,x,w) - \partial_3\sigma(t,x',w)| \leq C''_\sigma |x - x'|^{\gamma'_a},$$

for some $0 < \gamma'_a \leq 1$. Finally, the mapping $\varphi_2 : \mathbb{R}^d \ni x \mapsto \varphi_2(x) \in \mathbb{R}$ is supposed to be α_2-Hölder continuous, $0 < \alpha_2 \leq 1$.

This means that we let the diffusion coefficient a be only a Hölder-continuous function w.r.t. the space and law variable. We emphasize that (**HE**) implies (**H**$\mathcal{E}$).

In order to apply the Zvonkin's transformation, we do not need to solve the whole system of PDE. We can indeed regularize it and then exhibit a Lipschitz bound on the regularized solution and its space derivative depending only on known parameters in (**H**$\mathcal{E}$).

In our context, it is possible to mollify the coefficients b, a and the functions φ_i, $i = 1,2$, to obtain a sequence of smooth coefficients $(b_n, a_n)_{n\geq 1}$ (say bounded and infinitely differentiable with bounded derivatives of all order), and functions $(\varphi_1^n, \varphi_2^n)_{n\geq 1}$ (infinitely differentiable with bounded derivatives of all order greater than 1) that converges uniformly to b, a and φ_i, $i = 1,2$. If we denote by $\mathcal{A}^n$ the regularized version of the operator $\mathcal{A}$ and $\mathbf{u}^n$ the solution of the regularized version of (2.2), we have the following result:

Theorem 2.1. *For each n, the regularized system of PDEs* (2.2) *(*i.e. *with $\mathcal{A}^n$ and b_n instead of $\mathcal{A}$ and b) admits a unique classical solution $\mathbf{u}^n$, in the sense defined in* [CCD14]. *Moreover, there exists a positive $\mathcal{T}_{2.1}$, a positive constant $C_{2.1}$ and a positive number $\delta_{2.1}$ depending only on known parameters in* (**H**$\mathcal{E}$)*, such that, for all (t,x,μ) in $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and all n in $\mathbb{N}^*$ the mapping $\mathbf{u}^n$ satisfies:*

$$|\partial_\mu \mathbf{u}^n(t,x,\mu)(z)| + |\partial_\mu(\partial_x \mathbf{u}^n(t,x,\mu)(z))| + |\partial_x \mathbf{u}^n(t,x,\mu)| + \left|\partial_x^2 \mathbf{u}^n(t,x,\mu)\right| \leq C_{2.1} T^{\delta_{2.1}},$$

for all z in $\mathbb{R}^d$ and for T less than $\mathcal{T}_{2.1}$

The smoothness of the solution in space is not new. This phenomenon is well-known and follows from the ellipticity assumption assumed on a. What is more unexpected is that there are bounds obtained uniformly on the regularization procedure on the measure derivatives. Indeed, the coefficients of the PDE are not differentiable w.r.t. the argument μ and it is clear that any differentiation of $\mathbf{u}^n$ w.r.t. μ should involves the differentiation of the source term b w.r.t. this argument. So, by the chain rule, the bound should contain an estimate on the derivative of $\mathcal{P}_2(\mathbb{R}^d) \ni \nu \mapsto \langle \varphi_i^n, \nu \rangle$, which is given by $(\varphi_i^n)'$, so that this estimate should depend on the regularization procedure.

Nevertheless, it appears that for all $t < s$ in $[0,T]^2$ the derivative of the mapping $\mu \mapsto \langle \varphi_i, [X_s^{t,\mu}] \rangle$ can be estimated in terms of known parameters in (**H**$\mathcal{E}$) (in fact, combining an additional estimate on

$\partial_z\partial_\mu \mathbf{u}$ together with Arzelà-Ascoli Theorem, we are able to show that the estimate on $\partial_\mu \mathbf{u}$ holds for the mild solution of (2.2)).

We hence have a smoothing property in the measure space without any action of the Laplacian in that direction. This follows from the fact that the function φ_i is integrated against the law of the process so that there still is a Gaussian convolution of the initial data "μ" at any time $s > t$. Therefore, we recover the spatial smoothing.

To the best of our knowledge, this result is new, especially since we do not add any noise on the space of measures. This last aspect has been studied in [CDL14] where Mean Field Games with common noise are investigated. Roughly speaking, the Authors showed that common noise on the original system of interacting players translates into McKean-Vlasov system with random law (the family of probability measure of the underlying stochastic process having now a stochastic dynamic) which allows them to recover existence and uniqueness of Nash equilibrium.

Also, in the same spirit as us, David R. Baños studies in [Bn15] the Malliavin differentiability of processes having the same dynamic as (1.5) with Lipschitz coefficients. Although he does not consider explicitly a regularization phenomenon in the measure direction (the functions φ_i are continuously differentiables with bounded Lipschitz derivatives), he shows that the space regularization phenomenon still holds so that the mapping $x \mapsto \langle \phi, [X_s^{t,\delta_x}]\rangle \in \mathbb{R}$, ϕ in $\mathbb{L}_2([X_s^{t,\delta_x}])$ is weakly differentiable for any $s > t$ thanks to a stochastic perturbation approach of Bismut type.

Finally, let us emphasize that Remark 1 also applies for Theorem 2.1. It seems that the main structural assumption that allows to recover the spatial smoothing property in the measure direction comes from the particular "polynomial " dependence of the coefficients w.r.t. the measure.

3 Proof of the main result

We prove our main result by using a Picard's iteration. Let m be a positive integer, set $(X_t^\mu)^0 = X_0$ for all t in $[0,T]$ and define $(X_t^\mu)^{m+1}$ as the solution of:

$$(X_t^\mu)^{m+1} = X_0 + \int_0^t b(r,(X_r^\mu)^m, \langle \varphi_1, [(X_r^\mu)^m]\rangle)\mathrm{d}r + \int_0^t \sigma(r,(X_r^\mu)^m, \langle \varphi_2, [(X_r^\mu)^m]\rangle)\mathrm{d}B_r, \quad X_0 \sim \mu.$$

In order to remove the singular drift, we now have to use an Itô's Formula that matches our framework, *i.e.* stated on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$. This formula, involving chain rule for functions defined on $\mathcal{P}_2(\mathbb{R}^d)$, can be found in Section 3 of [CCD14]. By applying it to

$$(X_t^\mu)^{m+1} - \mathbf{u}^n(t,(X_t^\mu)^{m+1}, [(X_t^\mu)^{m+1}]),$$

we obtain that

$$\begin{aligned}(X_t^\mu)^{m+1} &= X_0 - \mathbf{u}^n(0,X_0,\mu) - \mathbf{u}^n(t,(X_t^\mu)^{m+1}, \left[(X_t^\mu)^{m+1}\right]) \\ &\quad + \int_0^t \sigma(s,(X_s^\mu)^m, \langle \varphi_2, [(X_s^\mu)^m]\rangle)\left[1 - \partial_2 \mathbf{u}^n(s,(X_s^\mu)^{m+1}, \left[(X_s^\mu)^{m+1}\right])\right]\mathrm{d}B_s + \mathcal{R}_t^m(n),\end{aligned} \tag{3.1}$$

where

$$\begin{aligned}\mathcal{R}_t^m(n) &= \int_0^t b_n(s,(X_s^\mu)^m, \langle \varphi_1^n, [(X_s^\mu)^m]\rangle) - b(s,(X_s^\mu)^m, \langle \varphi_1, [(X_s^\mu)^m]\rangle) \\ &\quad + (\mathcal{A}^n - \mathcal{A})\mathbf{u}^n(s,(X_s^\mu)^{m+1}, [(X_s^\mu)^{m+1}]))\mathrm{d}s.\end{aligned}$$

Hence,

$$\begin{aligned}|(X_t^\mu)^{m+1}-(X_t^\mu)^m| &\leq |\mathbf{u}^n(t,(X_t^\mu)^{m+1},\big[(X_t^\mu)^{m+1}\big])-\mathbf{u}^n(t,(X_t^\mu)^m,[(X_t^\mu)^m])|\\ &\quad+\Big|\int_0^t\Big\{\sigma(s,(X_r^\mu)^m,\langle\varphi_2,[(X_s^\mu)^m]\rangle)\big[1-\partial_2\mathbf{u}^n(s,(X_s^\mu)^{m+1},\big[(X_s^\mu)^{m+1}\big])\big]\\ &\qquad-\sigma(s,(X_s^\mu)^{m-1},\langle\varphi_2,[(X_s^\mu)^m]\rangle)\,[1-\partial_2\mathbf{u}^n(s,(X_s^\mu)^m,[(X_s^\mu)^m])]\Big\}\mathrm{d}B_s\Big|\\ &\quad+|\mathcal{R}_t^m(n)|+|\mathcal{R}_t^{m-1}(n)|.\end{aligned}$$

Let us now emphasize that when $T<\mathcal{T}_{2.1}$, Theorem 2.1 implies that for all (t,x) in $[0,T]\times\mathbb{R}^d$, all Z, Z' in $\mathbb{L}_2(\Omega)$:

$$\begin{aligned}&\Big|\partial_2^l\mathbf{u}^n(t,x,[Z])-\partial_2^l\mathbf{u}^n(t,x,[Z'])\Big|\\ &\leq\Big|\int_0^1\mathbb{E}\big[\partial_3\partial_2^l\mathbf{u}^n(t,x,[(1-\lambda)Z+\lambda Z'])(Z-Z')\big]\mathrm{d}\lambda\Big|\\ &\leq C_{2.1}T^{\delta_{2.1}}\mathbb{E}[|Z-Z'|].\end{aligned}$$

for $l=0,1$, so that for any T less than $\mathcal{T}_{2.1}$, there exists a positive δ, depending on known parameters in **(HE)**, such that:

$$\begin{aligned}\mathbb{E}\sup_{t\leq T}|(X_t^\mu)^{m+1}-(X_t^\mu)^m|^2 &\leq CT^\delta\mathbb{E}\sup_{t\leq T}|(X_t^\mu)^{m+1}-(X_t^\mu)^m|^2\\ &\quad+C\int_0^t\mathbb{E}\sup_{r\leq s}|(X_r^\mu)^m-(X_r^\mu)^{m-1}|^2\mathrm{d}s+2\mathbb{E}|\mathcal{R}_T^m(n)|^2+2\mathbb{E}|\mathcal{R}_T^{m-1}(n)|^2,\end{aligned}$$

Since for any m, $\mathbb{E}|\mathcal{R}_T^m(n)|^2$ tends uniformly to 0 as n tend to infinity, we can let n tend to infinity in the right hand side of the equation above and we get that

$$(1-CT^\delta)\mathbb{E}\sup_{t\leq T}|(X_t^\mu)^{m+1}-(X_t^\mu)^m|^2 \leq C_T\int_0^t\mathbb{E}\sup_{r\leq s}|(X_r^\mu)^m-(X_r^\mu)^{m-1}|^2\mathrm{d}s.$$

Finally, we can find a positive $\mathcal{T}$ depending on known parameters in (**HE**) such that for any T less than $\mathcal{T}$:

$$\mathbb{E}\sup_{t\leq T}|(X_t^\mu)^{m+1}-(X_t^\mu)^m|^2 \leq C_T\int_0^T\mathbb{E}\sup_{r\leq s}|(X_r^\mu)^m-(X_r^\mu)^{m-1}|^2\mathrm{d}s.$$

By induction, we deduce that for any T less than $\mathcal{T}$:

$$\mathbb{E}\sup_{t\leq T}|(X_t^\mu)^{m+1}-(X_t^\mu)^m|^2 \leq \frac{C_T^{m+1}T^m}{m!}\mathbb{E}\sup_{t\leq T}|(X_t^\mu)^1-(X_t^\mu)^0|^2\leq C'\frac{C_T^{m+1}T^m}{m!}.$$

So that $(X_t^\mu)^m$ converges almost surely to a solution X_t^μ of (1.5). We deduce the uniqueness part from the previous computations. We hence have existence and uniqueness of a solution on $[0,\mathcal{T}]$. We can then iterate the construction and obtain the result for all T in $\mathbb{R}^+$.

Remark 3. *For whom are aware with Zvonkin Transformation and regularization by noise phenomenon, asking the drift function b to be only a measurable and bounded function of the space variable (instead of the Hölder continuity assumed here) should be the right assumption. In that case, it seems that our strategy still apply. The main idea consists in showing that the usual ingredients used in that case still hold in the McKean-Vlasov setting. To do so, one may think about the law argument as a time parameter in* (4.4)*. Hence, one has to check that both the Krylov inequality and* $\mathbb{L}_p$ *estimate for the PDE solution still hold and then apply the Veretennikov arguments* [Ver80].

4 Estimation on the solution of the PDE: proof of Theorem 2.1

Notations. From now, we let $C, C', c, c', \tilde{C}, \tilde{C}', \dots$ be some positive constants depending only on known parameters in **(H$\mathcal{E}$)**.

Let us first reduce the problem. We emphasize that any component of the d-dimensional solution of the system of PDEs (2.2) above can be described by the solution of:

$$\begin{cases} (\partial_t + \mathcal{A})u(t,x,\mu) = \tilde{b}(t,x,\langle\varphi_1,\mu\rangle), \text{ on } [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \\ u(T,x,\mu) = 0, \end{cases} \tag{4.1}$$

where $\tilde{b} : \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ plays the role of one of the components of b. Hence, we only have to prove the estimates in Theorem 2.1 for the function u defined above. Next, we have that Theorem 2.1 is stated under regularized framework. For the sake of clarity, we forget the superscript n that follows from the regularization procedure in the following and we suppose that the following assumptions hold.

Assumptions (H$\mathcal{E}$**R**). We say that assumptions (H$\mathcal{E}$**R**) hold if assumptions (H$\mathcal{E}$) hold true and $b, \sigma, \varphi_1, \varphi_2$ are infinitely differentiable functions with bounded derivatives of all order, greater than 1 for the functions φ_1 and φ_2.

We know from [CCD14] that under **(H$\mathcal{E}$R)** this PDE admits a unique classical solution. Let us now give a suitable representation of this solution.

Under assumptions (H$\mathcal{E}$**R**) it follows from the Sznitman's note [Szn91] that equation (1.5) admits a unique strong solution. For any (t,μ) in $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$ and x in $\mathbb{R}^d$, its flow is the solution on $[t,T]$ of

$$X_s^{t,x,\mu} = x + \int_t^s b(r, X_r^{t,x,\mu}, \langle\varphi_1, [X_r^{t,\mu}]\rangle)\mathrm{d}r + \int_t^s \sigma(r, X_r^{t,x,\mu}, \langle\varphi_2, [X_r^{t,\mu}]\rangle)\mathrm{d}B_r. \tag{4.2}$$

Given the family of marginals $[X^{t,\mu}] := ([X_s^{t,\mu}])_{t\le s\le T}$ of the solution of (1.5), we can consider the stochastic system (4.2) as a linear system parametrized by the time dependent parameter $[X^{t,\mu}]$. We can then define for all (s',y') in $[0,T]\times\mathbb{R}^d$ the process $X^{s',y',[X^{t,\mu}]}$ as the solution of (4.2) on $[s',T]$ with starting point y' at time s' and whose coefficients depend on $[X^{t,\mu}]$ and we denote by $\mathcal{L}^{t,\mu}$ its generator.

It is then clear, thanks to the well posedness of (4.2) under (H$\mathcal{E}$**R**), that $X^{t,x,[X^{t,\mu}]} = X^{t,x,\mu}$. Finally, from classical theory of linear SDEs, the flow $X^{t,x,[X^{t,\mu}]}$ admits a transition density p which is also parametrized by $[X^{t,\mu}]$.

Since from the arguments of [CCD14] we have that, for all $(t,x,\mu) \in [0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R})$, the solution of the PDE (4.1) writes:

$$u(t,x,\mu) = \mathbb{E}\int_t^T \tilde{b}(s, X_s^{t,x,\mu}, \langle\varphi_1, [X_s^{t,\mu}]\rangle)\mathrm{d}s, \tag{4.3}$$

we deduce from the previous discussion that

$$u(t,x,\mu) = \int_t^T \int_{\mathbb{R}^d} \tilde{b}(s,y,\langle\varphi_1,[X_s^{t,\mu}]\rangle)p([X^{t,\mu}];t,x;s,y)\mathrm{d}y\mathrm{d}s. \tag{4.4}$$

In order to keep the notations clear, we only mention the dependence of p w.r.t. the initial data (t,μ) of (1.5) in the following and we forget its first argument when the starting time of (1.5) and (4.2) are the same. Hence, for all (s',y') and (s,y) in $[t,T]\times\mathbb{R}^d$: $p(t,\mu;s',y';s,y) := p([X^{t,\mu}];s',y;s,y)$ and $p(\mu;t,y';s,y) := p(t,\mu;t,y';s,y)$.

When differentiating the function u in the measure direction, we have to differentiate the integrand in the expression (4.4) in that direction. Thus, we have to estimate a quantity of the form

$$\partial_\mu\langle\varphi,[X_s^{t,\mu}]\rangle, \quad s\in(t,T].$$

Then, in the following, for any α-Hölder function ϕ, $\alpha\in(0,1]$, we denote by v the mapping:

$$v : (t<s,\mu,\phi) \in [0,T]^2\times\mathcal{P}_2(\mathbb{R}^d) \mapsto v_s(t,\mu,\phi) = \langle\phi,[X_s^{t,\mu}]\rangle, \tag{4.5}$$

and we prove in Section 5 the following Proposition.

Proposition 4.1. *Suppose that assumptions* (**H$\mathcal{ER}$**) *hold, let ϕ be some α-Hölder function from $\mathbb{R}^d$ to $\mathbb{R}$, let t in $[0,T]$ and let us denote by μ the law of the solution of* (1.5) *at time t. There exist a positive number $\mathcal{T}_{4.1}$ and a positive constant $C_{4.1}$, depending only on known parameters in* (**H$\mathcal{E}$**)*, such that for all z in $\mathbb{R}^d$ and s in $(t,T]$:*

$$|\partial_\mu v_s(t,\mu,\phi)(z)| \le C(s-t)^{(-1+\alpha)/2}, \tag{4.6}$$

for all T less than $\mathcal{T}_{4.1}$.

Moreover, we have from (4.4) that the derivative of u or $\partial_x u$ along the measure involves the derivative of the transition density p or $\partial_x p$ along the measure. We then have to obtain suitable control of these quantities. Here, these controls are summarized by the following Proposition whose proof is postponed to Section 6.

Proposition 4.2. *Suppose that assumptions* (**H$\mathcal{ER}$**) *hold, let t in $[0,T]$ and let us denote by μ the law of the solution of* (1.5) *at time t. Then, for all x in $\mathbb{R}^d$, for all (s,y) in $(t,T]\times\mathbb{R}^d$ and all z in $\mathbb{R}^d$: $p(\mu;t,x;s,y) \le \hat{p}_c(t,x;s,y)$ where $\hat{p}_c$ is the Gaussian like kernel defined by:*

$$\hat{p}_c(t,x;s,y) = \frac{c}{(s-t)^{d/2}}\exp\left(-c\frac{|y-x|^2}{(s-t)}\right), \tag{4.7}$$

where c depends on known parameters in (**H$\mathcal{E}$**) *only. Moreover, there exist two positive constants $C_{4.2}$ and $C'_{4.2}$, depending only on known parameters in* (**H$\mathcal{E}$**)*, such that for all x in $\mathbb{R}^d$, for all (s,y) in*

$(t,T]\times\mathbb{R}^d$ and all z in $\mathbb{R}^d$

$$\partial_\mu p(\mu;t,x;s,y)(z) \le C_{4.2}\sum_{i=1}^{2}(s-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s)(z)\hat{p}_c(t,x;s,y), \tag{4.8}$$

$$\partial_\mu\partial_x p(\mu;t,x;s,y)(z) \le C'_{4.2}\sum_{i=1}^{2}(s-t)^{\alpha_i/2-1}\overline{\partial_\mu v}^i_{\alpha_i/2-1}(\mu;t,s)(z)\hat{p}_c(t,x;s,y), \tag{4.9}$$

where we used the abusive notation

$$\overline{\partial_\mu v}^i_\gamma(\mu;t,s)(z) = \sup_{r\in[t,s]}\left\{(r-t)^{-\gamma}|\partial_\mu v_r(t,\mu,\varphi_i)(z)|\right\}.$$

We have now all the ingredients to complete the proof. Thanks to estimates (4.6) on $\partial_\mu v$ and (4.8) on $\partial_\mu p$ we deduce that we can invert the differentiation and integration operators when differentiating the right hand side of (4.4) w.r.t. the measure. Hence, the derivative of u in the measure direction writes, at any point z of $\mathbb{R}^d$:

$$\begin{aligned}\partial_\mu u(t,x,\mu)(z) &= \int_t^T\int_{\mathbb{R}^d}\partial_3\tilde{b}(s,y,\langle\varphi_1,[X_s^{t,\mu}]\rangle)\partial_\mu\langle\varphi_1,[X_s^{t,\mu}]\rangle(z)p(\mu;t,x;s,y)\mathrm{d}y\mathrm{d}s\\ &+ \int_t^T\int_{\mathbb{R}^d}\tilde{b}(s,y,\langle\varphi_1,[X_s^{t,\mu}]\rangle)\partial_\mu p(\mu;t,x;s,y)(z)\mathrm{d}y\mathrm{d}s,\end{aligned} \tag{4.10}$$

and satisfies, thanks to estimates (4.8) and (4.6)

$$\begin{aligned}&|\partial_\mu u(t,x,\mu)(z)|\\ &\le C\int_t^T\int_{\mathbb{R}^d}||\partial_3\tilde{b}||_\infty(s-t)^{(-1+\alpha)/2}p(\mu;t,x;s,y)\mathrm{d}y\mathrm{d}s\\ &\quad+\int_t^T\int_{\mathbb{R}^d}||\tilde{b}||_\infty C_{4.2}\sum_{i=1}^{2}(s-t)^{(-1+\alpha_i)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s)(z)\hat{p}_c(t,x;s,y)\mathrm{d}y\mathrm{d}s.\end{aligned}$$

Therefore we can deduce from (4.6) the there exist a positive constant C' and a positive number δ, depending only on known parameters in (**H**$\mathcal{E}$), such that:

$$|\partial_\mu u(t,x,\mu)(z)| \le C'T^\delta.$$

Now, we have:

$$\partial_x u(t,x,\mu) = \int_t^T\int_{\mathbb{R}^d}\tilde{b}(s,y,\langle\varphi_1,[X_s^{t,\mu}]\rangle)\partial_x p(\mu;t,x;s,y)\mathrm{d}y\mathrm{d}s. \tag{4.11}$$

Hence, we can differentiate the mapping $\partial_x u$ along the measure and by using the same arguments as above, with estimate (4.9) on $\partial_\mu\partial_x p$ instead of (4.8), we obtain that

$$|\partial_\mu(\partial_x u(t,x,\mu))(z)| \le C''T^{\delta'},$$

for some positive constant C'' and positive number δ', depending only on known parameters in (**H**$\mathcal{E}$).

Finally, the estimates on $\partial_x u$ and $\partial_x^2 u$ can be obtained by classical arguments, when viewing the argument μ as a parameter. See *e.g.* [Fri64]. This concludes the proof of Theorem 2.1.

5 Differentiation and estimation of v: proof of Proposition 4.1

With the notations defined in the previous section, we have that for all s in $(t,T]$,

$$\begin{aligned} v_s(t,\mu,\phi) &= \int_{\mathbb{R}^d}\phi(y)\int_{\mathbb{R}^d} p(\mu;t,x;s,y)\mathrm{d}\mu(x)\mathrm{d}y \\ &:= \int_{\mathbb{R}^d}\phi(y)P(t,\mu,\mu;s,y)\mathrm{d}y, \end{aligned}$$

where the function P is the function

$$P:(t<s,\lambda,\nu,y)\in[0,T]^2\times\mathcal{P}_2(\mathbb{R}^d)\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\mapsto P(t,(\lambda,\nu);s,y)=\int_{\mathbb{R}^d}p(\lambda;t,x;s,y)\mathrm{d}\nu(x).$$

With this notation and by using the fact that p is a density we have:

$$\begin{aligned} \partial_\mu v_s(t,\mu,\phi) &= \partial_\mu\int_{\mathbb{R}^d}\phi(y)P(t,\mu,\mu;s,x)\mathrm{d}y \\ &= \left[\partial_\lambda\int_{\mathbb{R}^d}\phi(y)P(t,\lambda,\mu;s,y)\mathrm{d}y\right]_{\lambda=\mu}+\left[\partial_\nu\int_{\mathbb{R}^d}\phi(y)P(t,\mu,\nu;s,y)\mathrm{d}y\right]_{\nu=\mu} \\ &= \left[\partial_\lambda\int_{\mathbb{R}^d}\phi(y)P(t,\lambda,\mu;s,y)\mathrm{d}y\right]_{\lambda=\mu}+\int_{\mathbb{R}^d}(\phi(y)-\phi(\xi))[\partial_\nu P(t,\mu,\nu;s,y)]_{\nu=\mu}\mathrm{d}y, \end{aligned}$$

whatever ξ in $\mathbb{R}^d$. Since by Fubini's Theorem we have

$$\begin{aligned} \partial_\lambda\int_{\mathbb{R}^d}\phi(y)P(t,\lambda,\mu;s,y)\mathrm{d}y &= \partial_\lambda\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\phi(y)-\phi(x))p(\lambda;t,x;s,y)\mathrm{d}\mu(x)\mathrm{d}y \\ &\quad+\partial_\lambda\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\phi(x)p(\lambda;t,x;s,y)\mathrm{d}\mu(x)\mathrm{d}y \\ &= \partial_\lambda\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\phi(y)-\phi(x))p(\lambda;t,x;s,y)\mathrm{d}\mu(x)\mathrm{d}y \\ &\quad+\partial_\lambda\int_{\mathbb{R}^d}\phi(x)\int_{\mathbb{R}^d}p(\lambda;t,x;s,y)\mathrm{d}y\mathrm{d}\mu(x) \\ &= \partial_\lambda\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\phi(y)-\phi(x))p(\lambda;t,x;s,y)\mathrm{d}\mu(x)\mathrm{d}y, \end{aligned}$$

and since by definition

$$\partial_\nu P(t,\mu,\nu;s,y)(\cdot) = \partial_x p(\mu;t,\cdot;s,y),$$

we deduce that for all ξ in $\mathbb{R}^d$, the derivative $\partial_\mu v$ taking at any point z in $\mathbb{R}^d$ writes

$$\begin{aligned}\partial_\mu v_s(t,\mu,\phi)(z) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\phi(y)-\phi(x))\partial_\mu p(\mu;t,x;s,y)(z)\mathrm{d}\mu(x)\mathrm{d}y \\ &\quad + \int_{\mathbb{R}^d}(\phi(y)-\phi(\xi))\partial_x p(\mu;t,z;s,y)\mathrm{d}y. \end{aligned} \tag{5.1}$$

So that for any given z in $\mathbb{R}^d$, by choosing $\xi = z$ we get

$$\begin{aligned}\partial_\mu v_s(t,\mu,\phi)(z) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\phi(y)-\phi(x))\partial_\mu p(\mu;t,x;s,y)(z)\mathrm{d}\mu(x)\mathrm{d}y \\ &\quad + \int_{\mathbb{R}^d}(\phi(y)-\phi(z))\partial_x p(\mu;t,z;s,y)\mathrm{d}y. \end{aligned}$$

Thanks to the estimates on the transition density p and its derivatives in the measure direction from Proposition 4.2, Fubini's Theorem, regularity assumed on ϕ and by using the Gaussian decay of $\hat{p}_c$ [2] we obtain the following bound:

$$\begin{aligned}|\partial_\mu v_s(t,\mu,\phi)(z)| &\le C\Bigg\{\sum_{i=1}^{2}(s-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s)(z)\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|y-x|^\alpha\hat{p}_c(t,x;s,y)\mathrm{d}\mu(x)\mathrm{d}y \\ &\quad + \int_{\mathbb{R}^d}|y-z|^\alpha(s-t)^{-1/2}\hat{p}_c(t,z;s,y)\mathrm{d}y\Bigg\} \\ &\le C'\Bigg\{\sum_{i=1}^{2}(s-t)^{(-1+\alpha_i+\alpha)/2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s)(z)+(s-t)^{(-1+\alpha)/2}\Bigg\}, \end{aligned} \tag{5.2}$$

which holds true for any α-Hölder function ϕ. Then, by choosing $\phi = \varphi_1$ (and so $\alpha = \alpha_1$), by multiplying both sides by $(s-t)^{(1-\alpha_1)/2}$, we deduce from a circular argument that there exists a positive time $\mathcal{T}'$ depending only on known parameters in $(\mathbf{H}\mathcal{E})$, such that for all T less than $\mathcal{T}'$:

$$(s-t)^{(1-\alpha_1)/2}\,|\partial_\mu v_s(t,\mu,\varphi_1)(z)| \le C''\Bigg\{(s-t)^{\alpha_2/2}\overline{\partial_\mu v}^{2}_{(-1+\alpha_2)/2}(\mu;t,s)(z)+1\Bigg\}. \tag{5.3}$$

By plugging this estimate in (5.2) and by iterating this argument (choosing $\phi = \varphi_2$ so that $\alpha = \alpha_2$, then multiplying both sides by $t^{(1-\alpha_2/2)}$ and using a circular argument) we obtain that there exists a positive time $\mathcal{T}''$, depending only on known parameters in $(\mathbf{H}\mathcal{E})$, such that for all T less than $\mathcal{T}''$:

$$(s-t)^{(1-\alpha_2)/2}\,|\partial_\mu v_s(t,\mu,\varphi_2)(z)| \le C''. \tag{5.4}$$

Again, by plugging this estimate in (5.3) and then using the resulting estimate together with (5.4) in (5.2), we finally deduce that there exists a positive time $\mathcal{T}$, depending only on known parameters in $(\mathbf{H}\mathcal{E})$, such that for all T less than $\mathcal{T}$:

$$(s-t)^{(1-\alpha)/2}\,|\partial_\mu v_s(t,\mu,\phi)(z)| \le C''',$$

which concludes the proof.

[2] *i.e.* the inequality: $\forall\eta>0,\ \forall q>0,\ \exists\bar{C}>0$ s.t. $\forall\sigma>0,\ \sigma^q e^{-\eta\sigma}\le\bar{C}$.

6 Estimation of the transition density p

This section is dedicated to the proof of Proposition 4.2 and so, to the study of the transition density p of the flow (4.2). Under (**H$\mathcal{E}$R**) it is clear that for all initial data (t,μ) in $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, equation (1.5) admits a unique solution $X^{t,\mu}$. Thus, we can suppose that the family of probability measures $[X^{t,\mu}]$ acts as a time dependent parameter in (4.2), so that the unique solution of (4.2) has a classical transition density p parametrized by the family of probability measures $[X^{t,\mu}]$. Once the law dependence is fixed, we can now use a classical parametrix expansion of McKean-Singer type [MS67] for linear processes in order to represent the transition density p.

The parametrix expansion of Mckean and Singer is based on the following observation: in small time, the transition density of a (smooth enough) process should be closed enough to the transition density of the associated frozen process (*i.e.* whose coefficients are constants and fixed at the (final) value of the process). Hence, the transition density of interest can be expanded in terms of the frozen transition density. Since the frozen transition density usually enjoys well known properties, *e.g.* it has an explicit form or can be estimated by explicit (and nice) functions, this expansion allows to estimate the original transition density. As a consequence, this method requires a good knowledge of the frozen transition density and on the associated frozen process.

Thus, this section is organized as follows: we first introduce in subsection 6.1 the frozen process and its associated transition density and give its explicit expression. Then, we give the estimates on the frozen transition density and its derivative. In subsection 6.2 we show how the transition density of (4.2) can be expanded in terms of the transition density of the frozen process. Hence, we obtain an explicit expression of the transition density p of (4.2) which can be estimated. These estimations are done in subsection 6.3 and lead to proof of Proposition 4.2. Finally, we suppose throughout this section that $T<1$.

6.1 The frozen system

Let (t,μ) in $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, for any point ξ in $\mathbb{R}^d$, we define the frozen flow as the solution of:

$$\tilde{X}_s^{s',y',\xi,[X^{t,\mu}]} = y' + \int_{s'}^{s} b(r,\xi,\langle\varphi_1,[X_r^{t,\mu}]\rangle)\mathrm{d}r + \int_{s'}^{s}\sigma(r,\xi,\langle\varphi_2,[X_r^{t,\mu}]\rangle)\mathrm{d}B_r, \tag{6.1}$$

for all (s',y') in $[t,T]\times\mathbb{R}^d$. Under (**H$\mathcal{E}$R**), it is clear that this flow exists and is unique, moreover, it has a transition density $\tilde{p}$ defined for all (s,y) in $[t,T]\times\mathbb{R}^d$ by:

$$\tilde{p}^{\xi}(t,\mu;s',y';s,y) = \frac{1}{(2\pi)^{d/2}}\big[\det[a^{\xi}_{s',s}(t,\mu)]\big]^{-1/2}\exp\left(-\frac{1}{2}\Big|\big[a^{\xi}_{s',s}(t,\mu)\big]^{-1/2}\big(y-y'-m^{\xi}_{s',s}(t,\mu)\big)^*\Big|^2\right), \tag{6.2}$$

where we adopted the same convention of notations for $\tilde{p}$ as for p and where

$$m^{\xi}_{s',s}(t,\mu) = \int_{s'}^{s} b(r,\xi,\langle\varphi_1,[X_r^{t,\mu}]\rangle)\mathrm{d}r, \tag{6.3}$$

$$a^{\xi}_{s',s}(t,\mu) = \int_{s'}^{s} a(r,\xi,\langle\varphi_2,[X_r^{t,\mu}]\rangle)\mathrm{d}r. \tag{6.4}$$

The frozen transition density (6.2) admits Gaussian type bounds, namely, we prove in Appendix A the following result.

Proposition 6.1. *Suppose that hypothethis* (**H$\mathcal{E}$R**) *holds. Let t in $[0,T]$ and let μ denotes the law of the process* (1.5) *at time t. Then:*

- *there exists a positive constant $\bar{C}_{6.1}$, depending only on known parameters in* (**H$\mathcal{E}$**)*, such that*

$$\forall s < s' \in [0,T]^2,\ \forall y, y' \in \mathbb{R}^d,\ \forall \xi \in \mathbb{R}^d, \quad \tilde{p}^{\xi}(t,\mu;s',y';s,y) \le \bar{C}_{6.1}\hat{p}_c(s',y';s,y), \tag{6.5}$$

  where $\hat{p}_c$ is the Gaussian like kernel defined by (4.7)*;*
- *there exist two positive constants $C_{6.1}$ and $C'_{6.1}$, depending on known parameters in* (**H$\mathcal{E}$**) *only, such that for all s in $(t,T]$, for all ξ in $\mathbb{R}^d$ and all x, y in $\mathbb{R}^d$:*

$$\left|\partial_\mu \tilde{p}^{\xi}(\mu;t,x;s,y)(z)\right| \le C_{6.1}\sum_{i=1}^{2}(s-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(t,x;s,y), \tag{6.6}$$

$$\left|\partial_\mu \partial_x \tilde{p}^{\xi}(\mu;t,x;s,y)(z)\right| \le C'_{6.1}\sum_{i=1}^{2}(s-t)^{\alpha_i/2-1}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s)(z)\hat{p}_c(t,x;s,y), \tag{6.7}$$

  for all z in $\mathbb{R}^d$;
- *for all $s' < s$ in $(t,T]^2$, for all y', y in $\mathbb{R}^d$:*

$$\left|\partial_\mu \tilde{p}^{\xi}(t,\mu;s',y';s,y)(z)\right| \le C\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s)(z)\hat{p}_c(s',y';s,y),$$

$$\left|\partial_\mu \partial_{y'} \tilde{p}^{\xi}(t,\mu;s',y';s,y)(z)\right| \le C'(s-s')^{-1/2}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(s',y';s,y),$$

$$\left|\partial_\mu \partial^2_{y'} \tilde{p}^{\xi}(t,\mu;s',y';s,y)(z)\right| \le C''(s-s')^{-1}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(s',y';s,y),$$

  for all z in $\mathbb{R}^d$.

Finally, we have that the generator of the frozen flow (6.1) is given by

$$\tilde{\mathcal{L}}^{\xi,t,\mu}_{s',y'} := b(s',\xi,\langle\varphi_1,[X^{t,\mu}_{s'}]\rangle)\partial_{y'} + \frac{1}{2}\mathrm{Tr}\left[a(s',\xi,\langle\varphi_2,[X^{t,\mu}_{s'}]\rangle)\partial^2_{y'}\right]. \tag{6.8}$$

Above, the subscript (s',y') means that the coefficients of the operator are evaluated at time s' and that the differentiation operator acts on the space variable y'.

6.2 The parametrix expansion

We now give the parametrix representation of the transition density p of (4.2). Proof of such a result is classical, we nevertheless wrote it since it allows to understand the crucial estimates, which will be useful in the sequel.

Proposition 6.2. *There exists a smoothing kernel $H : [0,T]\times\mathcal{P}_2(\mathbb{R}^d)\times[0,T]\times\mathbb{R}^d\times[0,T]\times\mathbb{R}^d \to \mathbb{R}$ given by*

$$H(t,\mu;s',y';s,y) = \left(\tilde{\mathcal{L}}^{y,t,\mu} - \mathcal{L}^{t,\mu}\right)_{s',y'}\tilde{p}^{y}(t,\mu;s',y';s,y), \tag{6.9}$$

where we recall that $\mathcal{L}^{t,\mu}$ is the generator of (4.2)*, such that the transition density p of the flow X defined by* (4.2) *writes:*

$$p(t,\mu;s',y';s,y)=\tilde{p}^{y}(t,\mu;s',y';s,y)+\sum_{k=1}^{+\infty}\int_{s'}^{s}\int_{\mathbb{R}^d}H^{\otimes k}(t,\mu;r,u;s,y)\tilde{p}^{y'}(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r, \quad (6.10)$$

where $H^{\otimes k}$ is recursively defined by:

$$H^{\otimes k+1}(t,\mu;s',y';s,y)=\int_{s'}^{s}\int_{\mathbb{R}^d}H^{\otimes k}(t,\mu;r,u;s,y)H(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r, \quad (6.11)$$

and $H^{\otimes 0}=\mathrm{Id}$.

Proof. Let (s,y) belong to $[0,T]\times\mathbb{R}^d$, the transition density $\tilde{p}^{y}(t,\mu;\cdot,\cdot;s,y)$ satisfies the Fokker-Planck equation:

$$\begin{cases}\partial_{s'}\tilde{p}^{y}(t,\mu;s',y';s,y)+\tilde{\mathcal{L}}^{y,t,\mu}_{s',y'}\tilde{p}^{y}(t,\mu;s',y';s,y)=0, \quad (s',y')\in[0,s)\times\mathbb{R}^d,\\ \tilde{p}^{y}(t,\mu;s,y';s,y)=\delta_y(y'),\end{cases}$$

which can be rewritten as

$$\begin{cases}\partial_{s'}\tilde{p}^{y}(t,\mu;s',y';s,y)+\mathcal{L}^{t,\mu}_{s',y'}\tilde{p}^{y}(t,\mu;s',y';s,y)=(\mathcal{L}^{t,\mu}-\tilde{\mathcal{L}}^{y,t,\mu})_{s',y'}\tilde{p}^{y}(t,\mu;s',y';s,y), \quad (s',y')\in[0,s)\times\mathbb{R}^d\\ \tilde{p}^{y}(t,\mu;s',y';s,y)=\delta_y(y').\end{cases}$$

Note that $p(t,\mu;\cdot,\cdot;s,y)$ is a fundamental solution of this PDE. Therefore $\tilde{p}^{y}(t,\mu;\cdot,\cdot;s,y)$ writes, for all $(s',y')\in[0,s]\times\mathbb{R}^d$:

$$\tilde{p}^{y}(t,\mu;s',y';s,y)=p(t,\mu;s',y';s,y)+\int_{s'}^{s}\int_{\mathbb{R}^d}(\mathcal{L}^{t,\mu}-\tilde{\mathcal{L}}^{y,t,\mu})_{r,u}\tilde{p}^{y}(t,\mu;r,u;s,y)p(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r.$$

Hence, by iterating N times this procedure, we obtain that

$$\begin{aligned}p(t,\mu;s'y';s,y) &= \tilde{p}^{y}(t,\mu;s',y';s,y)+\sum_{k=1}^{N}\int_{s'}^{s}\int_{\mathbb{R}^d}H^{\otimes k}(t,\mu;r,u;s,y)\tilde{p}^{y'}(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r\\ &\quad+\int_{s'}^{s}\int_{\mathbb{R}^d}H^{\otimes N+1}(t,\mu;r,u;s,y)p(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r. \quad (6.12)\end{aligned}$$

In order to obtain the parametrix expansion of p, depending only on known quantities (*i.e.* on the smoothing kernel H defined by (6.9) and on the transition density of the frozen process $\tilde{p}$) the idea consists in letting N tend to infinity. To this aim, we need a "good" estimate on the approximation error. These controls are the estimate (6.5) in Proposition 6.1 and the following Lemma.

Lemma 6.3. *Under assumption* (**H$\mathcal{E}$R**) *the following assertion holds: there exists a positive constant $C_k^{6.3}$ given by:*

$$C_k^{6.3}=C_{6.3}^{k}\prod_{r=1}^{k-1}\beta(\gamma_a r/2,\gamma_a/2), \quad (6.13)$$

where $C_{6.3}$ is a positive constant depending only on known parameters in (**H$\mathcal{E}$**)*, β denotes the bêta-function and with the convention $\prod_1^0\equiv 1$, such that for all $s'<s$ in $[t,T]^2$ and y',y in $\mathbb{R}^d$:*

$$|H^{\otimes k}(t,\mu;s',y';s,y)|\leq C_k^{6.3}(s-s')^{k\gamma_a/2-1}\hat{p}_c(s',y';s,y).$$

On the one hand, the first term in the right hand side of (6.12) is controlled by a convolution of two Gaussian functions which is still Gaussian and it is clear from the asymptotic properties of the beta-function (that are recalled in Appendix B) that the series converges. On the other hand, the N^{th} convolution of the kernel H tends uniformly to 0 as N tends to infinity (recall that T is small) and since p is a density, we deduce that the second term in the right hand side of (6.12) tends to 0. Therefore, the density p writes:

$$\begin{aligned} p(\mu;t,x;s,y) &= \tilde{p}^{y}(\mu;t,x;s,y) \\ &\quad + \sum_{k\geq 1}\int_t^s\int_{\mathbb{R}^d} H^{\otimes k}(t,\mu;s',y';s,y)\tilde{p}^{y'}(\mu;t,x;s',y')\mathrm{d}y'\mathrm{d}s'. \end{aligned} \tag{6.14}$$

□

Proof of Lemma 6.3. By using classical parametrix arguments (see Chapter 1 of [Fri64]) and by the definition (6.9) of H, we deduce that there exist two positive constants $C_{6.3}$ and c depending only on known parameters in ($\mathbf{H}\mathcal{E}$) such that for all $s' < s \in [t,T]^2$ and $y', y \in \mathbb{R}^d$:

$$|H(t,\mu;s',y';s,y)| \leq C_{6.3}(s-s')^{\gamma_a/2-1}\hat{p}_c(s',y';s,y). \tag{6.15}$$

Suppose now as an induction hypothesis that for all $s' < s \in [t,T]^2$ and $y' \in \mathbb{R}^d$:

$$|H^{\otimes k}(t,\mu;s',y';s,y)| \leq C_k^{6.3}(s-s')^{k\gamma_a/2-1}\hat{p}_c(s',y';s,y), \tag{6.16}$$

where $C_k^{6.3}$ is defined by (6.13). Recall that for all integer k, $H^{\otimes k+1}$ is recursively defined by

$$H^{\otimes k+1}(t,\mu;s',y';s,y) = \int_{s'}^s\int_{\mathbb{R}^d} H^{\otimes k}(t,\mu;r,u;s,y)H(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r. \tag{6.17}$$

Hence, by plugging (6.15) and (6.16) in (6.17) and using the Gaussian convolution we obtain that:

$$|H^{\otimes k+1}(t,\mu;s',y';s,y)| \leq C_k^{6.3}C_{6.3}\int_{s'}^s (s-r)^{\gamma_a k/2-1}(r-s')^{\gamma_a/2-1}\mathrm{d}r\hat{p}_c(s',y';s,y),$$

and by the change of variable $r = (s-s')r' + s'$ we have:

$$\begin{aligned} |H^{\otimes k+1}(t,\mu;s',y';s,y)| &\leq C_{6.3}C_k^{6.3}(s-s')^{(k+1)\gamma_a/2-1}\int_0^1 (1-r')^{\bar{\gamma}_a k/2-1}(r')^{\gamma_a/2-1}\mathrm{d}r'\hat{p}_c(s',y';s,y) \\ &= C_{6.3}C_k^{6.3}(s-s')^{(k+1)\gamma_a/2-1}\beta(\gamma_a k/2,\gamma_a/2)\hat{p}_c(s',y';s,y) \\ &= C_{k+1}^{6.3}(s-s')^{(k+1)\gamma_a/2-1}\hat{p}_c(s',y';s,y). \end{aligned}$$

□

6.3 Differentiation and estimation of the density p along the measure

We are now ready to prove Proposition 4.2. We have that

$$p(\mu;t,x;s,y) = \tilde{p}^{y}(\mu;t,x;s,y) + \sum_{k=1}^{+\infty}\int_t^s\int_{\mathbb{R}^d} H^{\otimes k}(t,\mu;s',y';s,y)\tilde{p}^{y'}(\mu;t,x;s',y')\mathrm{d}y'\mathrm{d}s',$$

so that the derivative of p w.r.t. μ writes, at any point z in $\mathbb{R}^d$:

$$\begin{aligned}\partial_\mu p(\mu;t,x;s,y)(z) &= \partial_\mu \tilde{p}^y(\mu;t,x;s,y)(z) \\ &\quad +\partial_\mu\left(\sum_{k=1}^{+\infty}\int_t^s\int_{\mathbb{R}^d} H^{\otimes k}(t,\mu;s',y';s,y)\tilde{p}^{y'}(t,\mu;t,x;s',y')\mathrm{d}y'\mathrm{d}s'\right)(z).\end{aligned} \tag{6.18}$$

Then, in order to invert the integration and differentiation operators in the right hand side of the above equation, we have to show that for all z in $\mathbb{R}^d$,

$$\begin{aligned}(s',y') &\mapsto \partial_\mu H^{\otimes k}(t,\mu;s',y';s,y)(z)\tilde{p}^{y'}(\mu;t,x;s',y') \\ &\quad +H^{\otimes k}(t,\mu;s',y';s,y)\partial_\mu\tilde{p}^{y'}(\mu;t,x;s',y')(z),\end{aligned} \tag{6.19}$$

is suitably bounded. More precisely, we have to obtain a Gaussian control on the derivative of the kth iteration of the smoothing kernel $\partial_\mu H^{\otimes k}$ and on $\partial_\mu\tilde{p}$ so that the parametrix expansion still holds, in the same spirit of the proof of Proposition 6.2. These controls are given by the estimates on the frozen transition density in Proposition 6.1 and the following Lemma.

Lemma 6.4. *Let t in $[0,T]$ and let μ be the law of the process (1.5) at time t. For all positive integer k, there exists a positive constant $\tilde{C}_k^{6.4}$, depending only on known parameter in $(\mathbf{H}\mathcal{E})$ and recursively defined by:*

$$\tilde{C}_k^{6.4} = (C_{k-1}^{6.3}\tilde{C} + C\tilde{C}_{k-1}^{6.4})\beta((k-1)\bar{\gamma}_a/2,\bar{\gamma}_a/2),$$

for $k\geq 2$ and $\tilde{C}_1^{6.4}=\tilde{C}$ and where $\bar{\gamma}_a=\gamma_a\wedge\gamma_a'$ such that

$$\begin{aligned}&\left|\partial_\mu H^{\otimes k}(t,\mu;s',y';s,y)(z)\right| \\ &\leq \tilde{C}_k^{6.4}(s'-s)^{k\bar{\gamma}_a/2-1}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(s',y';s,y),\end{aligned}$$

for all $s'<s\in(t,T]^2$, y, y' in $\mathbb{R}^d$ and z in $\mathbb{R}^d$.

From Lemma 6.4 and estimate (6.5) in Proposition 6.1 we have that for all k:

$$\begin{aligned}&\left|\partial_\mu H^{\otimes k}(t,\mu;s',y';s,y)(z)\tilde{p}^{y'}(\mu;t,x;s',y')\right| \\ \leq\ & \tilde{C}_k^{6.4}\bar{C}_{6.1}(s-s')^{k\bar{\gamma}_a/2-1}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(t,x;s,'y')\hat{p}_c(s',y';s,y),\end{aligned}$$

and from estimate (6.6) of Proposition 6.1 and Lemma 6.3 we have that for all k

$$\begin{aligned}&\left|H^{\otimes k}(t,\mu;s',y';s,y)\partial_\mu\tilde{p}^{y'}(\mu;t,x;s',y')(z)\right| \\ \leq\ & C_k^{6.3}C_{6.1}(s-s')^{k\bar{\gamma}_a/2-1}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(t,x;s',y')\hat{p}_c(s',y';s,y).\end{aligned}$$

We can hence invert the differentiation and integration operators in the second term in the right hand side of (6.18) and using property of Gaussian convolution we get that there exists a positive constant C, depending on known parameters in $(\mathbf{H}\mathcal{E})$ only, such that

$$\begin{aligned}
&\partial_\mu p(\mu;t,x;s,y)(z)\\
&\le C(s-t)^{(\alpha_i-1)/2}\left(C_{6.1}+\sum_{k\ge 1}\left\{(C_k^{6.3}C_{6.1}+\tilde{C}_k^{6.4}\bar{C}_{6.1})\beta(k\bar{\gamma_a}/2,(\alpha_i+1)/2)(s-t)^{k\bar{\gamma}_a/2}\right\}\right)\\
&\quad\times\hat{p}_c(t,x;s,y)\sum_{i=1}^{2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s)(z),
\end{aligned}$$

so that estimate (4.8) of Proposition 4.2 follows from the estimates on the parametrix constants $\tilde{C}_k^{6.4}$ and on the beta-function given in Appendix B.

For the second assertion it is well seen from usual parametrix technique that

$$\partial_x p(\mu;t,x;s,y)=\partial_x\tilde{p}^y(\mu;t,x;s,y)+\sum_{k=1}^{+\infty}\int_t^s\int_{\mathbb{R}^d}H^{\otimes k}(t,\mu;s',y';s,y)\partial_x\tilde{p}^{y'}(\mu;t,x;s',y')\mathrm{d}y'\mathrm{d}s',$$

so that the derivative of $\partial_x p$ w.r.t. μ writes, at any point z in $\mathbb{R}^d$:

$$\begin{aligned}
\partial_\mu(\partial_x p)(\mu;t,x;s,y)(z) &= \partial_\mu(\partial_x\tilde{p}^y)(\mu;t,x;s,y)(z)\\
&\quad+\partial_\mu\left(\sum_{k=1}^{+\infty}\int_t^s\int_{\mathbb{R}^d}H^{\otimes k}(t,\mu;s',y';s,y)(\partial_x\tilde{p}^{y'})(\mu;t,x;s',y')\mathrm{d}y'\mathrm{d}s'\right)(z).
\end{aligned}$$

We can use the same arguments as above with estimates (6.7) instead of (6.6) in Proposition 6.1 and we obtain that

$$\begin{aligned}
&\partial_\mu\partial_x p(\mu;t,x;s,y)(z)\\
&\le C(s-t)^{\alpha_i/2-1}\left(C'_{6.1}+\sum_{k\ge 1}\left\{(C_k^{6.3}C'_{6.1}+\tilde{C}_k^{6.4}\bar{C}_{6.1})\beta(k\bar{\gamma_a}/2,(\alpha_i+1)/2)(s-t)^{k\bar{\gamma}_a/2}\right\}\right)\\
&\quad\times\hat{p}_c(t,x;s,y)\sum_{i=1}^{2}\overline{\partial_\mu v}^{i}_{(\alpha_i-1)/2}(\mu;t,s)(z),
\end{aligned}$$

from which we deduce estimate (4.9) of Proposition 4.2.

A Proofs of Lemmas 6.4 and 6.1

In order to avoid heavy notations, the proofs are done in the real case ($d=1$). We also recall that $T<1$.

Proof of Lemma 6.4. Recall that by definition

$$\begin{aligned}
H(t,\mu;s',y';s,y) &= \left(b(s',y,v_{s'}(t,\mu,\varphi_1))-b(s',y',v_{s'}(t,\mu,\varphi_1))\right)\partial_x\tilde{p}(\mu;s',y';s,y)\\
&\quad+\frac{1}{2}\left[(a(s',y,v_{s'}(t,\mu,\varphi_2))-a(s,y',v_{s'}(t,\mu,\varphi_2)))\partial_x^2\tilde{p}(\mu;s',y';s,y)\right],
\end{aligned}$$

so that, for any z in $\mathbb{R}^d$:

$$
\begin{aligned}
&\partial_\mu H(t,\mu;s',y';s,y)(z)\\
&=(b(s',y,v_{s'}(t,\mu,\varphi_1))-b(s',y',v_{s'}(t,\mu,\varphi_1)))\partial_\mu\partial_x\tilde{p}(\mu;s',y';s,y)(z)\\
&\quad+(\partial_3 b(s',y,v_{s'}(t,\mu,\varphi_1))-\partial_3 b(s',y',v_{s'}(t,\mu,\varphi_1)))\partial_\mu v_{s'}(t,\mu,\varphi_1)(z)\partial_x\tilde{p}(\mu;s',y';s,y)\\
&\quad+\frac{1}{2}(a(s',y,v_{s'}(t,\mu,\varphi_2))-a(s',y',v_{s'}(t,\mu,\varphi_2)))\partial_\mu\partial_x^2\tilde{p}(\mu;s',y';s,y)(z)\\
&\quad+\frac{1}{2}(\partial_3 a(s',y,v_s(t,\mu,\varphi_2))-\partial_3 a(s',y',v_{s'}(t,\mu,\varphi_2)))\partial_\mu v_{s'}(t,\mu,\varphi_2)(z)\partial_x^2\tilde{p}(\mu;s',y';s,y).
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\left|\partial_\mu H(t,\mu;s',y';s,y)(z)\right|\\
&\le C\Bigg\{||b||_\infty\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{\,i}_{(\alpha_i-1)/2}(\mu;t,s')(z)(s-s')^{-1/2}\\
&+||\partial_3 b||_\infty(s'-t)^{(\alpha_1-1)/2}\overline{\partial_\mu v}^{\,1}_{(\alpha_1-1)/2}(\mu;t,s')(z)(s-s')^{-1/2}\\
&+||a||_{\bar\gamma_a}|y-y'|^{\bar\gamma_a}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{\,i}_{(\alpha_i-1)/2}(\mu;t,s')(z)(s-s')^{-1}\\
&+||\partial_3 a||_{\gamma'_a}|y-y'|^{\gamma'_a}(s'-t)^{(\alpha_2-1)/2}\overline{\partial_\mu v}^{\,2}_{(\alpha_2-1)/2}(\mu;t,s')(z)(s-s')^{-1}\Bigg\}\hat{p}_c(s',y';s,y),
\end{aligned}
$$

where $\bar\gamma_a=\gamma_a\wedge\gamma'_a$. Therefore, by using the Gaussian decay of $\hat{p}_c$:

$$
\begin{aligned}
&\left|\partial_\mu H(t,\mu;s',y';s,y)(z)\right|\\
&\le C\Bigg\{||b||_\infty\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\sum_{i=1}^{2}\overline{\partial_\mu v}^{\,i}_{(\alpha_i-1)/2}(\mu;t,s')(z)(s-s')^{-1/2}\\
&+||\partial_3 b||_\infty(s'-t)^{(\alpha_1-1)/2}\overline{\partial_\mu v}^{\,1}_{(\alpha_1-1)/2}(\mu;t,s')(z)(s-s')^{-1/2}\\
&+||a||_{\bar\gamma_a}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{\,i}_{(\alpha_i-1)/2}(\mu;t,s')(z)(s-s')^{-1+\bar\gamma_a/2}\\
&+||\partial_3 a||_{\gamma'_a}(s'-t)^{(\alpha_2-1)/2}\overline{\partial_\mu v}^{\,2}_{(\alpha_2-1)/2}(\mu;t,s')(z)(s-s')^{-1+\gamma'_a/2}\Bigg\}\hat{p}_c(s',y';s,y).
\end{aligned}
$$

So, there exists a positive constant $\tilde{C}$ depending on known parameters in **(H$\mathcal{E}$)** such that:

$$
\begin{aligned}
&\left|\partial_\mu H(t,\mu;s',y';s,y)(z)\right|\\
&\le\tilde{C}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^{\,i}_{(\alpha_i-1)/2}(\mu;t,s')(z)(s-s')^{\bar\gamma_a/2-1}\hat{p}_c(s',y';s,y).
\end{aligned}
\tag{A.1}
$$

Assume now as an induction hypothesis that for all (s',y') in $(t;s)\times\mathbb{R}^d$:

$$\left|\partial_\mu H^{\otimes k}(t,\mu;s',y';s,y)(z)\right| \le \tilde{C}_k^{6.4}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z)(s-s')^{k\bar{\gamma}_a/2-1}\hat{p}_c(s',y';s,y),$$

where

$$\tilde{C}_k^{6.4} = (C_{k-1}^{6.4}\tilde{C} + C\tilde{C}_{k-1}^{6.4})\beta((k-1)\bar{\gamma}_a/2,\bar{\gamma}_a/2).$$

We then have

$$\begin{aligned}
&\partial_\mu H^{\otimes k+1}(t,\mu;s',y';s,y)(z) \qquad (A.2)\\
&= \int_{s'}^{s}\int_{\mathbb{R}^d}\partial_\mu H^{\otimes k}(t,\mu;r,u;s,y)(z)H(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r\\
&\quad + \int_{s'}^{s}\int_{\mathbb{R}^d} H^{\otimes k}(t,\mu;r,u;s,y)\partial_\mu H(t,\mu;s',y';r,u)(z)\mathrm{d}u\mathrm{d}r.
\end{aligned}$$

We can bound the first term in the right hand side by using the induction hypothesis above, the estimation (6.17) on H and the property of the Gaussian convolution:

$$\begin{aligned}
&\left|\int_{s'}^{s}\int_{\mathbb{R}^d}\partial_\mu H^{\otimes k}(t,\mu;r,u;s,y)(z)H(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r\right|\\
&\le C_{6.3}\tilde{C}_k^{6.4}\int_{s'}^{s}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,r)(z)(s-r)^{k\bar{\gamma}_a/2-1}(r-s')^{\bar{\gamma}_a/2-1}\mathrm{d}r\hat{p}_c(s',y';s,y)\\
&\le C_{6.3}\tilde{C}_k^{6.4}\sum_{i=1}^{2}(s'-t)^{(\alpha_i-1)/2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(s',y';s,y)\int_{s'}^{s}(s-r)^{k\bar{\gamma}_a/2-1}(r-s')^{\bar{\gamma}_a/2-1}\mathrm{d}r.
\end{aligned}$$

By the change of variable $r=(s-s')r'+s'$, one can show that

$$\int_{s'}^{s}(s-r)^{k\bar{\gamma}_a/2-1}(r-s')^{\bar{\gamma}_a/2-1}\mathrm{d}r \le (s-s')^{(k+1)\bar{\gamma}_a/2-1}\beta(k\bar{\gamma}_a/2,\bar{\gamma}_a/2),$$

so that

$$\begin{aligned}
&\left|\int_{s'}^{s}\int_{\mathbb{R}^d}\partial_\mu H^{\otimes k}(t,\mu;r,u;s,y)(z)H(t,\mu;s',y';r,u)\mathrm{d}u\mathrm{d}r\right| \qquad (A.3)\\
&\le C_{6.3}\tilde{C}_k^{6.4}\sum_{i=1}^{2}\overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z)\hat{p}_c(s',y';s,y)(s-s')^{(k+1)\bar{\gamma}_a/2-1}\beta(k\bar{\gamma}_a/2,\bar{\gamma}_a/2).
\end{aligned}$$

Now, we bound the second term in the right hand side of (A.2). We have from Lemma 6.3:

$$\left|H^{\otimes k}(t,\mu;r,u;s,y)\right| \le C_k^{6.3}(s-r)^{k\bar{\gamma}_a/2-1}\hat{p}_c(r,u;s,y),$$

where

$$C_k^{6.3} = C_{6.3}^k \prod_{l=1}^{k-1} \beta(r\bar{\gamma}_a/2, \bar{\gamma}_a/2).$$

So that, thanks to estimate (A.1):

$$\begin{aligned} &\left| \int_{s'}^{s} \int_{\mathbb{R}^d} H^{\otimes k}(t,\mu;r,u;s,y) \partial_\mu H(t,\mu;s',y';r,u)(z) \mathrm{d}u \mathrm{d}r \right| \\ &\quad \leq \sum_{i=1}^{2} (s'-t)^{(\alpha_i-1)/2} \overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z) C_k^{6.3} \tilde{C} \int_{s'}^{s} (s-r)^{k\bar{\gamma}_a/2-1} (r-s')^{\bar{\gamma}_a/2-1} \mathrm{d}r. \end{aligned} \tag{A.4}$$

Hence, by plugging (A.3) and (A.4) in (A.2) we get:

$$\begin{aligned} &\left| \partial_\mu H^{\otimes k+1}(t,\mu;s',y';s,y)(z) \right| \\ &\leq (s-s')^{(k+1)\bar{\gamma}_a/2-1} \sum_{i=1}^{2} (s'-t)^{(\alpha_i-1)/2} \overline{\partial_\mu v}^i_{(\alpha_i-1)/2}(\mu;t,s')(z) (C_k^{6.3}\tilde{C} + C\tilde{C}_k^{6.4}) \beta(k\bar{\gamma}_a/2, \bar{\gamma}_a/2), \end{aligned}$$

and the induction is true since:

$$\tilde{C}_{k+1}^{6.4} = (C_k^{6.3}\tilde{C} + C\tilde{C}_k^{6.4}) \beta(k\bar{\gamma}_a/2, \bar{\gamma}_a/2).$$

□

Proof of Lemma 6.1. We begin with the following Claim.

Claim A.1. *The following estimates hold:*

- $\left|\partial_\mu m^\xi_{t,s}(t,\mu)(z)\right| \leq (s-t)^{(\alpha_1+1)/2} ||\partial_3 b||_\infty \overline{\partial_\mu v}^1_{(1-\alpha_1)/2}(\mu;t,s)(z)$,
- $\left|\partial_\mu a^\xi_{t,s}(t,\mu)(z)\right| \leq (s-t)^{(\alpha_2+1)/2} ||\partial_3 a||_\infty \overline{\partial_\mu v}^2_{(1-\alpha_2)/2}(\mu;t,s)(z)$,

and for all $s' \in (t,T]$:

- $\left|\partial_\mu m^\xi_{s',s}(t,\mu)(z)\right| \leq (s-s')(s'-t)^{(\alpha_1-1)/2} ||\partial_3 b||_\infty \overline{\partial_\mu v}^1_{(1-\alpha_1)/2}(\mu;t,s)(z)$,
- $\left|\partial_\mu a^\xi_{t,s}(t,\mu)(z)\right| \leq (s-s')(s'-t)^{(\alpha_2-1)/2} ||\partial_3 a||_\infty \overline{\partial_\mu v}^2_{(1-\alpha_2)/2}(\mu;t,s)(z)$.

We also recall the classical estimate coming from the uniform ellipticity of a: there exists $\Lambda > 0$ such that, for all positive γ:

$$\frac{1}{\left[a^\xi_{s',s}(t,\mu)\right]^\gamma} < \Lambda^{-\gamma}(s-s')^{-\gamma}. \tag{A.5}$$

The derivative of $\tilde{p}$ evaluated at any point z in $\mathbb{R}$ is then given by:

$$\begin{aligned}
&\partial_\mu \tilde{p}^\xi(t,\mu;s',y';s,y)(z)\\
&= \Bigg(-\frac{1}{2}\frac{\partial_\mu a^\xi_{s',s}(t,\mu)(z)}{\big[a^\xi_{s',s}(t,\mu)\big]} + \Bigg(\frac{1}{2}\frac{\big(\partial_\mu a^\xi_{s',s}(t,\mu)(z)\big)\big(y-y'-m^\xi_{s',s}(t,\mu)\big)}{\big[a^\xi_{s',s}(t,\mu)\big]^{3/2}} + \frac{\partial_\mu m^\xi_{s',s}(t,\mu)(z)}{\big[a^\xi_{s',s}(t,\mu)\big]}\Bigg)\\
&\quad\times\left(\frac{y-y'-m^\xi_{s',s}(t,\mu)}{\big[a^\xi_{t,s}([X^{t,\mu}])\big]^{1/2}}\right)\Bigg)\frac{1}{\sqrt{2\pi}}\frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}}\exp\left(-\frac{1}{2}\left|\big[a^\xi_{s',s}(t,\mu)\big]^{-1/2}\big(y-y'-m^\xi_{s',s}(t,\mu)\big)\right|^2\right).
\end{aligned}$$

Now, by using the Gaussian decay of $\tilde{p}$, and estimates of Claim A.1, we obtain that, when $s'=t$:

$$\begin{aligned}
&\big|\partial_\mu \tilde{p}^\xi(t,\mu;t,y';s,y)(z)\big|\\
&\le C\Bigg(\Lambda^{-1}||\partial_3 a||_\infty \sup_{r\in[t,s]}\overline{\partial_\mu v}^2_{(1-\alpha_2)/2}(\mu;t,s)(z)(s-t)^{(\alpha_2-1)/2}\\
&\quad+\Lambda^{-1}||\partial_3 b||_\infty \sup_{r\in[t,s]}\overline{\partial_\mu v}^1_{(1-\alpha_1)/2}(\mu;t,s)(z)(s-t)^{(\alpha_1-1)/2}\Bigg)\\
&\quad\frac{1}{\sqrt{2\pi}}\frac{1}{\big[a^\xi_{t,s}(t,\mu)\big]^{1/2}}\exp\left(-c\frac{1}{2}\left|\big[a^\xi_{t,s}(t,\mu)\big]^{-1/2}\big(y-y'-m^\xi_{t,s}(t,\mu)\big)\right|^2\right),
\end{aligned}$$

for some positive constant C and c, with c strictly less than 1. We now compute the space derivatives: for all $s'<s$ in $[t,T]^2$

$$\begin{aligned}
&\partial_x \tilde{p}^\xi(t,\mu;s',y';s,y)\\
&= \frac{\big(y-y'-m^\xi_{s',s}(t,\mu)\big)}{\big[a^\xi_{s',s}(t,\mu)\big]}\frac{1}{\sqrt{2\pi}}\frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}}\exp\left(-c\frac{1}{2}\left|\big[a^\xi_{s',s}(t,\mu)\big]^{-1/2}\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^*\right|^2\right),
\end{aligned}$$

which gives

$$|\partial_x \tilde{p}^\xi(t,\mu;s',y';s,y)| \le C(s-t)^{-1/2}\frac{1}{\sqrt{2\pi}}\frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}}\exp\left(-\frac{1}{2}\left|\big[a^\xi_{s',s}(t,\mu)\big]^{-1/2}\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^*\right|^2\right).$$

Next

$$\begin{aligned}
\partial_x^2 \tilde{p}^\xi(t,\mu;s',x;s,y) &= \left(\frac{\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^2}{\big[a^\xi_{s',s}(t,\mu)\big]^2}-\frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]}\right)\\
&\quad\times\frac{1}{\sqrt{2\pi}}\frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}}\exp\left(-\frac{1}{2}\left|\big[a^\xi_{s',s}(t,\mu)\big]^{-1/2}\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^*\right|^2\right),
\end{aligned}$$

so that

$$\partial_x^2\tilde{p}^\xi(t,\mu;s',y';s,y)$$
$$\leq C(1+\Lambda^{-1})(s-s')^{-1}\frac{1}{2\pi}\frac{1}{\left[a^\xi_{s',s}(t,\mu)\right]^{1/2}}\exp\left(-\frac{1}{2}\left|\left[a^\xi_{s',s}(t,\mu)\right]^{-1/2}\left(y-y'-m^\xi_{s',s}(t,\mu)\right)^*\right|^2\right).$$

Concerning the cross derivatives, we have, at any point z of $\mathbb{R}$:

$$\partial_\mu\partial_x\tilde{p}^\xi(t,\mu;s',y';s,y)(z)$$
$$=\left\{-\frac{\left(\partial_\mu m^\xi_{s',s}(t,\mu)(z)\right)}{\left[a^\xi_{s',s}(t,\mu)\right]}-\frac{3}{2}\frac{\left(\partial_\mu a^\xi_{s',s}(t,\mu)(z)\right)\left(y-y'-m^\xi_{s',s}(t,\mu)\right)}{\left[a^\xi_{s',s}(t,\mu)\right]^2}\right.$$
$$-\frac{1}{2}\frac{\partial_\mu a^\xi_{s',s}(t,\mu)(z)}{\left[a^\xi_{s',s}(t,\mu)\right]}+\left(\frac{1}{2}\frac{\left(\partial_\mu a^\xi_{s',s}(t,\mu)(z)\right)\left(y-x-m^\xi_{t,s}([X^{t,\mu}])\right)}{\left[a^\xi_{t,s}([X^{t,\mu}])\right]^{3/2}}+\frac{\partial_\mu m^\xi_{s',s}(t,\mu)(z)}{\left[a^\xi_{s',s}(t,\mu)\right]^{1/2}}\right)$$
$$\left.\times\left(\frac{y-y'-m^\xi_{s',s}(t,\mu)}{\left[a^\xi_{s',s}(t,\mu)\right]^{1/2}}\right)\right)\Bigg\}$$
$$\frac{1}{\sqrt{2\pi}}\frac{1}{\left[a^\xi_{s',s}(t,\mu)\right]^{1/2}}\exp\left(-\frac{1}{2}\left|\left[a^\xi_{s',s}(t,\mu)\right]^{-1/2}\left(y-y'-m^\xi_{s',s}(t,\mu)\right)^*\right|^2\right).$$

Thus we have: when $s'=t$,

$$|\partial_\mu\partial_x\tilde{p}^\xi(t,\mu;t,y';s,y)(z)|$$
$$\leq\left(\left((s-t)^{(\alpha_1-1)/2}+(s-t)^{\alpha_1}\right)||\partial_3 b||_\infty\overline{\partial_\mu v}^1_{(1-\alpha_1)/2}(\mu;t,s)(z)\right.$$
$$\left.+\left((s-t)^{\alpha_2/2-1}+2(s-t)^{(\alpha_2-1)/2}\right)||\partial_3 a||_\infty\overline{\partial_\mu v}^2_{(1-\alpha_2)/2}(\mu;t,s)(z)\right)$$
$$\times\frac{C\Lambda^{-1}}{\left[a^\xi_{t,s}([X^{t,\mu}])\right]^{1/2}}\exp\left(-c\frac{1}{2}\left|\left[a^\xi_{t,s}([X^{t,\mu}])\right]^{-1/2}\left(y-y'-m^\xi_{t,s}([X^{t,\mu}])\right)^*\right|^2\right),$$

and when $s'>t$,

$$|\partial_\mu\partial_x\tilde{p}^\xi(t,\mu;s',y';s,y)(z)|$$
$$\leq\left(\left((s'-t)^{(\alpha_1-1)/2}+(s'-t)^{(\alpha_1-1)/2}\right)||\partial_3 b||_\infty\overline{\partial_\mu v}^1_{(1-\alpha_1)/2}(\mu;t,s)(z)\right.$$
$$\left.+\left((s'-t)^{(\alpha_2-1)/2}+2(s'-t)^{(\alpha_2-1)/2}\right)||\partial_3 a||_\infty\overline{\partial_\mu v}^2_{(\alpha_2-1)/2}(\mu;t,s)(z)\right)$$
$$\times\frac{C(s-s')^{-(1/2)}}{\left[a^\xi_{s',s}(t,\mu)\right]^{1/2}}\exp\left(-c\frac{1}{2}\left|\left[a^\xi_{s',s}(t,\mu)\right]^{-1/2}\left(y-y'-m^\xi_{s',s}(t,\mu)\right)^*\right|^2\right).$$

We conclude with:

$$
\begin{aligned}
&\partial_\mu \partial_x^2 \tilde{p}^\xi(t,\mu;s',y';s,y)(z)\\
&= \Bigg\{ \Bigg(\frac{2\partial_\mu m^\xi_{s',s}([X^{t,\mu}])(z)(y-y'-m^\xi_{s',s}([X^{t,\mu}]))}{\big[a^\xi_{s',s}(t,\mu)\big]^2} - \frac{2\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^2 \partial_\mu a^\xi_{s',s}(t,\mu)(z)}{\big[a^\xi_{s',s}(t,\mu)\big]^3}\\
&+ \frac{\partial_\mu a^\xi_{s',s}(t,\mu)(z)}{\big[a^\xi_{s',s}(t,\mu)\big]^2} \Bigg) + \left(\frac{\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^2}{\big[a^\xi_{s',s}(t,\mu)\big]^{3/2}} - \frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}} \right) \left(-\frac{\partial_\mu a^\xi_{s',s}(t,\mu)(z)}{2\big[a^\xi_{s',s}(t,\mu)\big]^{3/2}} \right)\\
&+ \left(\frac{\big(y-y'-m^\xi_{s',s}(t,\mu)\big)^2}{\big[a^\xi_{s',s}(t,\mu)\big]^2} - \frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]} \right)\\
&\quad \times \Bigg(-\frac{1}{2}\frac{\partial_\mu a^\xi_{s',s}(t,\mu)(z)}{\big[a^\xi_{s',s}(t,\mu)\big]} + \left(\frac{1}{2}\frac{\big(\partial_\mu a^\xi_{s',s}(t,\mu)(z)\big)\big(y-y'-m^\xi_{s',s}(t,\mu)\big)}{\big[a^\xi_{s',s}(t,\mu)\big]^{3/2}} + \frac{\partial_\mu m^\xi_{s',s}(t,\mu)(z)}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}} \right)\\
&\qquad \times \left(\frac{y-y'-m^\xi_{s',s}(t,\mu)}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}} \right) \Bigg) \Bigg\}\\
&\times \frac{1}{2\pi} \frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}} \exp\left(-\frac{1}{2} \Big| \big[a^\xi_{s',s}(t,\mu)\big]^{-1/2} \big(y-y'-m^\xi_{s',s}(t,\mu)\big)^* \Big|^2 \right).
\end{aligned}
$$

Which gives, for all $s' > t$,

$$
\begin{aligned}
&\Big| \partial_\mu \partial_x^2 \tilde{p}^\xi(\mu;t,x;s,y)(z) \Big|\\
&\le \Bigg(||\partial_3 b||_\infty (s-s')^{-1/2} (s'-t)^{(\alpha_1-1)/2} \overline{\partial_\mu v}^1_{(1-\alpha_1)/2}(\mu;t,s)(z)\\
&+ ||\partial_3 a||_\infty \Big((s-s')^{-1} + (s-s')^{-1/2} \Big) (s'-t)^{(\alpha_2-1)/2} \overline{\partial_\mu v}^2_{(1-\alpha_2)/2}(\mu;t,s)(z) \Bigg)\\
&\times \frac{1}{\sqrt{2\pi}} \frac{1}{\big[a^\xi_{s',s}(t,\mu)\big]^{1/2}} \exp\left(-\frac{1}{2} \Big| \big[a^\xi_{s',s}(t,\mu)\big]^{-1/2} \big(y-y'-m^\xi_{s',s}(t,\mu)\big)^* \Big|^2 \right).
\end{aligned}
$$

□

Proof of Claim A.1. From chain rule we have, for all ξ in $\mathbb{R}$ and for all $s' < s \in [t,T]$:

$$
\begin{aligned}
\partial_\mu m^\xi_{s',s}(t,\mu)(z) &= \Big(\partial_\mu \int_{s'}^s b(r,\xi,\langle \varphi_1, [X^{t,\mu}_r)]\rangle) \mathrm{d}r \Big)(z)\\
&= \int_{s'}^s \partial_3 b(r,\xi,\langle \varphi_1, [X^{t,\mu}_r)]\rangle) \partial_\mu v_r(t,\mu,\varphi_1)(z) \mathrm{d}r,
\end{aligned}
$$

for all z in $\mathbb{R}$ and where we recall that v is defined by (4.5). So,

$$
\big| \partial_\mu m^\xi_{s',s}(t,\mu)(z) \big| \le ||\partial_3 b||_\infty \sup_{r\in[t,s]} \big\{ (r-t)^{(1-\alpha_1)/2} |\partial_\mu v_r(t,\mu,\varphi_1)(z)| \big\} \int_{s'}^s (r-t)^{(\alpha_1-1)/2} \mathrm{d}r.
$$

Hence, for all z in $\mathbb{R}$

$$\left|\partial_\mu m^\xi_{s',s}(t,\mu)(z)\right| \leq ||\partial_3 b||_\infty \sup_{r\in[t,s]} \left\{(r-t)^{(1-\alpha_1)/2}|\partial_\mu v_r(t,\mu,\varphi_1)(z)|\right\}(s-t)^{(\alpha_1+1)/2},$$

when $s'=t$ and

$$\left|\partial_\mu m^\xi_{s',s}(t,\mu)(z)\right| \leq ||\partial_3 b||_\infty \sup_{r\in[t,s]} \left\{(r-t)^{(1-\alpha_1)/2}|\partial_\mu v_r(t,\mu,\varphi_1)(z)|\right\}(s'-s)(s'-t)^{(\alpha_1-1)/2},$$

when $s'>t$. Next we have,

$$\begin{aligned}\partial_\mu a^\xi_{s',s}(t,\mu)(z) &= \left(\partial_\mu \int_{s'}^s a(r,\xi,\langle\varphi_2,[X^{t,\mu}_r)]\rangle)\mathrm{d}r\right)(z)\\ &= \int_{s'}^s \partial_3 a(r,\xi,\langle\varphi_2,[X^{t,\mu}_r)]\rangle)\partial_\mu v_r(t,\mu,\varphi_2)(z)\mathrm{d}r,\end{aligned}$$

for all z in $\mathbb{R}$. Therefore

$$\left|\partial_\mu a^\xi_{s',s}(t,\mu)(z)\right| \leq ||\partial_3 a||_\infty \sup_{r\in[t,s]} \left\{(r-t)^{(1-\alpha_2)/2}|\partial_\mu v_r(t,\mu,\varphi_2)(z)|\right\}\int_t^s (r-t)^{(\alpha_2-1)/2}\mathrm{d}r.$$

So

$$\left|\partial_\mu a^\xi_{s',s}(t,\mu)(z)\right| \leq ||\partial_3 a||_\infty \sup_{r\in[t,s]} \left\{(r-t)^{(1-\alpha_2)/2}|\partial_\mu v_r(t,\mu,\varphi_2)(z)|\right\}(s-t)^{(\alpha_2+1)/2},$$

when $s'=t$ and

$$\left|\partial_\mu a^\xi_{s',s}(t,\mu)(z)\right| \leq ||\partial_3 a||_\infty \sup_{r\in[t,s]} \left\{(r-t)^{(1-\alpha_2)/2}|\partial_\mu v_r(t,\mu,\varphi_2)(z)|\right\}(s'-s)(s'-t)^{(\alpha_2-1)/2},$$

when $s'>t$. □

B Asymptotic properties of the parametrix constants

Claim B.1. *There exists a strictly finite and strictly positive integer $K(\gamma_a)$ such that for all $k \geq K(\gamma_a)$:*

$$C^{6.3}_k = C(K(\gamma_a))C^k_{6.3}\frac{4^k}{\gamma_a^k(k!)^{\gamma_a/2}},$$

where $C(K(\gamma_a)) = 4^{-K(\gamma_a)}(K(\gamma_a)!)^{\gamma_a/2}\prod_{l=1}^{K(\gamma_a)-1}\beta(l\gamma_a/2,\gamma_a/2)$.

Proof. Let $K(\gamma_a)$ be equal to $\lceil 2/\gamma_a\rceil$, so that for all $k \geq K(\gamma_a)$

$$k\gamma_a/2 - 1 \geq 0,$$

and recall that

$$\beta(\gamma_a k/2,\gamma_a/2) = \int_0^1 (1-s)^{k\gamma_a/2}s^{\gamma_a/2-1}\mathrm{d}s.$$

For all positive ϵ strictly less than 1, we have, for all $k \geq K(\gamma_a)$:

$$\begin{aligned}\int_0^1 (1-s)^{\gamma_a k/2-1} s^{\gamma_a/2-1}\mathrm{d}s &= \int_0^1 (1-s)^{\gamma_a/2-1} s^{\gamma_a k/2-1}\mathrm{d}s \qquad \text{(B.1)}\\ &= \int_0^{1-\epsilon} (1-s)^{\gamma_a/2-1} s^{\gamma_a k/2-1}\mathrm{d}s + \int_{1-\epsilon}^1 (1-s)^{\gamma_a/2-1} s^{\gamma_a k/2-1}\mathrm{d}u \\ &\leq \frac{1}{\epsilon^{1-\gamma_a/2}} \int_0^{1-\epsilon} s^{\gamma_a k/2-1}\mathrm{d}s + \int_{1-\epsilon}^1 (1-s)^{\gamma_a/2-1}\mathrm{d}s \\ &\leq \frac{1}{\epsilon^{1-\gamma_a/2}} \frac{2}{k\gamma_a} + \frac{2}{\gamma_a}\epsilon^{\gamma_a/2}.\end{aligned}$$

So that, by letting $\epsilon = 1/k$ we have

$$\int_0^1 (1-s)^{k/2-1} s^{-1/2}\mathrm{d}s \leq \frac{4}{\gamma_a k^{\gamma_a/2}}, \qquad \text{(B.2)}$$

which gives the desired result. □

Claim B.2. *There exists a strictly finite and strictly positive integer $K := K(\gamma_a, \gamma_a')$ such that for all $k \geq K$:*

$$\tilde{C}_{k+1}^{6.4} = (C_k^{6.3}\tilde{C} + C\tilde{C}_k^{6.4})\beta(k\bar{\gamma}_a/2, \bar{\gamma}_a/2) \leq \tilde{C}(\gamma_a, \gamma_a')\left((k-K)\frac{\kappa^k}{\bar{\gamma}_a^k (k!)^{\bar{\gamma}_a/2}} + \frac{(4C)^{k+1}}{\bar{\gamma}_a^k (k!)^{\bar{\gamma}_a/2}}\right),$$

for some positive real $\kappa \geq 4C_{6.3}$ and some positive constant $\tilde{C}(\gamma_a, \gamma_a')$.

Proof. Let $K(\bar{\gamma}_a)$ be equal to $\lceil 2/\bar{\gamma}_a \rceil$, so that for all $k \geq K(\bar{\gamma}_a)$

$$k\bar{\gamma}_a/2 - 1 \geq 0.$$

Then, let $k \geq K := K(\bar{\gamma}_a) \vee K(\gamma_a)$. We know from the proof of Claim B.1 that:

$$\beta(\bar{\gamma}_a k/2, \bar{\gamma}_a/2) \leq \frac{4}{\bar{\gamma}_a k^{\bar{\gamma}_a/2}}.$$

Hence, from definition of $C_k^{6.3}$ we have that:

$$\tilde{C}_{k+1}^{6.4} \leq C(K(\gamma_a))\frac{\tilde{C}\kappa^k}{\bar{\gamma}_a^k (k!)^{\bar{\gamma}_a/2}} + \tilde{C}_k^{6.4} C \frac{4}{\bar{\gamma}_a k^{\bar{\gamma}_a/2}},$$

so that

$$\frac{\tilde{C}_{k+1}^{6.4}}{(k+1)^{\bar{\gamma}_a/2}} \leq C(K(\gamma_a))\frac{\tilde{C}\kappa^k}{\bar{\gamma}_a^k ((k+1)!)^{\bar{\gamma}_a/2}} + \frac{\tilde{C}_k^{6.4}}{k^{\bar{\gamma}_a/2}} \frac{4C}{\bar{\gamma}_a (k+1)^{\bar{\gamma}_a/2}}.$$

Let us define

$$A_k = \frac{\tilde{C}_k^{6.4}}{k^{\gamma_a/2}}, \quad M_{k+1} = \frac{C(K(\gamma_a))\tilde{C}\kappa^k}{\bar{\gamma}_a^k ((k+1)!)^{\bar{\gamma}_a/2}}, \quad D_{k+1} = \frac{4C}{\bar{\gamma}_a ((k+1)!)^{\bar{\gamma}_a/2}}.$$

Hence, we have that

$$A_{k+1} \leq M_{k+1} + A_k D_{k+1}.$$

Since κ is such that $\kappa \geq 4C_{6.3}$ we obtain:

$$M_k D_{k+1} = \frac{4C_{6.3}}{\gamma_a((k+1)!)^{\bar{\gamma}_a/2}} \frac{C(K(\gamma_a))\tilde{C}\kappa^{k-1}}{\gamma_a^{k-1}(k!)^{\bar{\gamma}_a/2}} \leq \frac{\kappa^k \tilde{C} C(K(\bar{\gamma}_a))}{\bar{\gamma}_a^k((k+1)!)^{\bar{\gamma}_a/2}} = M_{k+1}.$$

Therefore, by induction

$$A_{k+1} \leq (k-K)M_{k+1} + \prod_{i=K+1}^{k+1} \frac{4C_{6.3}}{\bar{\gamma}_a i^{\bar{\gamma}_a/2}},$$

which implies that

$$\frac{\tilde{C}_{k+1}}{(k+1)^{\bar{\gamma}_a/2}} \leq (k-K)\frac{C(K(\gamma_a))\tilde{C}\kappa^k}{\bar{\gamma}_a^k(k+1)!)^{\bar{\gamma}_a/2}} + \frac{(4C)^{k+1}}{\bar{\gamma}_a^k((k+1)!)^{\bar{\gamma}_a/2}} C(K),$$

where $C(K) = (4C)^{-K}\bar{\gamma}_a^K K!$. Then,

$$\tilde{C}_{k+1} \leq (k-K)C(K(\gamma_a))\tilde{C}\frac{\kappa^k}{\bar{\gamma}_a^k(k!)^{\bar{\gamma}_a/2}} + C(K)\frac{(4C)^{k+1}}{\bar{\gamma}_a^k(k!)^{\bar{\gamma}_a/2}}.$$

□

References

[BLPR14] Rainer Buckdahn, Juan Li, Shige Peng, and Catherine Rainer. Mean-field stochastic differential equations and associated PDEs. *arXiv:1407.1215 [math]*, July 2014. arXiv: 1407.1215.

[Bn15] David R. Baños. The Bismut-Elworthy-Li formula for mean-field stochastic differential equations. *arXiv:1510.06961 [math]*, October 2015. arXiv: 1510.06961.

[Car10] Pierre Cardaliaguet. Notes on mean field games. Technical report, 2010.

[CCD14] Jean-François Chassagneux, Dan Crisan, and François Delarue. A Probabilistic approach to classical solutions of the master equation for large population equilibria. *arXiv:1411.3009 [math]*, November 2014. arXiv: 1411.3009.

[CDL14] Rene Carmona, Francois Delarue, and Daniel Lacker. Mean field games with common noise. *arXiv:1407.6181 [math]*, July 2014. arXiv: 1407.6181.

[DL89] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Inventiones Mathematicae*, 98(3):511–547, 1989.

[Fla11] Franco Flandoli. *Random perturbation of PDEs and fluid dynamic models*, volume 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.

[Fri64] Avner Friedman. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[HMC06] Minyi Huang, Roland P. Malhamé, and Peter E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.

[KR05] Nicolai Vladimirovitch Krylov and Michael Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probability Theory and Related Fields*, 131(2):154–196, February 2005.

[LL06a] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I. Le cas stationnaire. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 343(9):619–625, 2006.

[LL06b] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *Comptes Rendus Mathématique. Académie des Sciences. Paris*, 343(10):679–684, 2006.

[LL07] J.M. Lasry and P.L. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[MS67] Henry P. McKean, Jr. and Isadore Singer. Curvature and the eigenvalues of the Laplacian. *Journal of Differential Geometry*, 1(1):43–69, 1967.

[Szn91] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.

[Ver80] Alexander Ju. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Matematicheski\u\ Sbornik. Novaya Seriya*, 111(153)(3):434–452, 480, 1980.

[Zha11] Xicheng Zhang. Stochastic Homeomorphism Flows of SDEs with Singular Drifts and Sobolev Diffusion Coefficients. *Electronic Journal of Probability*, 16(0), June 2011.

[Zvo74] Alexander Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

Univ. Savoie Mont Blanc, CNRS, LAMA, F-73000 Chambéry, France

E-mail address: pe.deraynal@univ-smb.fr