WELL-POSEDNESS FOR SOME NON-LINEAR DIFFUSION PROCESSES AND RELATED PDE ON THE WASSERSTEIN SPACE

Paul-Eric Chaudru de Raynal, Noufel Frikha

WELL-POSEDNESS FOR SOME NON-LINEAR DIFFUSION PROCESSES AND RELATED PDE ON THE WASSERSTEIN SPACE

PAUL-ERIC CHAUDRU DE RAYNAL AND NOUFEL FRIKHA

Abstract. In this paper, we investigate the well-posedness of the martingale problem for non-linear stochastic differential equations (SDEs) in the sense of McKean-Vlasov under mild assumptions on the coefficients as well as classical solutions for a class of associated linear partial differential equations (PDEs) defined on $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, for any $T>0$, $\mathcal{P}_2(\mathbb{R}^d)$ being the Wasserstein space, that is, the space of probability measures on $\mathbb{R}^d$ with a finite second-order moment. The martingale problem is addressed by a perturbation argument on $\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, for non-linear coefficients including any bounded continuous drift and diffusion coefficient satisfying some structural assumption in the measure sense that covers a large class of interaction. Some new well-posedness results in the strong sense also directly stem from the previous analysis. Under additional assumptions, we then establish the existence and smoothness of the associated density as well as Gaussian type bounds, the derivatives with respect to the measure being understood in the sense introduced by P.-L. Lions. Finally, existence and uniqueness for the related linear Cauchy problem with irregular terminal condition and source term among the considered class of non-linear interaction is addressed.

Contents

1. Introduction

In this work, we are interested in some non-linear stochastic differential equations (SDEs for short) in the sense of McKean-Vlasov with dynamics:

$$X_t^\xi = \xi + \int_0^t b(s, X_s^\xi, [X_s^\xi])ds + \int_0^t \sigma(s, X_s^\xi, [X_s^\xi])dW_s, \quad [\xi]\in\mathcal{P}_2(\mathbb{R}^d), \tag{1.1}$$

driven by a q-dimensional $W=(W^1,\cdots,W^q)$ Brownian motion with coefficients $b:\mathbb{R}_+\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d$ and $\sigma:\mathbb{R}_+\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d\otimes\mathbb{R}^q$. Here and throughout, we denote by $[\theta]$ the law of the random variable θ. This kind of dynamics are also referred to as mean-field or McKean-Vlasov SDEs as it describes the limiting behaviour of an individual particle evolving within a large system of particles interacting through its empirical measure, as the size of the population grows to infinity. More generally, the behaviour of the particle system is ruled by the so-called propagation of chaos phenomenon as originally studied by McKean [McK67] and then investigated by Sznitman [Szn91]. Roughly speaking, it says that if the initial conditions of a finite subset of the original system of particles become independent of each other, as the size of the whole system grows to infinity, then the dynamics of the particles of the finite subset synchronize and also become independent. Since the original works of Kac [Kac56] in kinetic theory and of McKean [McK66] in non-linear parabolic partial differential equations (PDEs for short), many authors have investigated theoretical and numerical aspects of McKean-Vlasov SDEs

Date: November 16, 2018.
2000 *Mathematics Subject Classification.* Primary 60H10, 93E03; Secondary 60H30, 35K40.
Key words and phrases. McKean-Vlasov SDEs; weak uniqueness; martingale problem; parametrix expansion; density estimates.

under various settings such as: well-posedness of related martingale problem, propagation of chaos and other limit theorems, probabilistic representations to non-linear parabolic PDEs and their numerical approximation schemes. We refer to Tanaka [Tan78], Gärtner [Gär88], [Szn91] among others.

Well-posedness in the weak or strong sense of McKean-Vlasov SDEs have been intensively investigated under various settings by many authors during the last decades, see e.g. Funaki [Fun84], Oelschläger [Oel84], [Gär88], [Szn91], Jourdain [Jou97], and more recently, Li and Min [LM16], Chaudru de Raynal [CdR15], Mishura and Veretenikov [MV18], Lacker [Lac18] and Hammersley et al. [HvS18] for a short sample.

We here revisit the problem of the unique solvability of the SDE (1.1) by tackling the corresponding formulation of the martingale problem. The main idea relies on Stroock & Varadhan's [SV79] perturbation argument applied on a suitable space. More precisely, we follow the methodology originally proposed by Bass and Perkins [BP09] in the framework of non-degenerate, non-divergence and time-homogeneous diffusion operators under the assumption that the diffusion matrix is a bounded and Hölder continuous function. We also refer the reader to Bass and Perkins [BP03], Menozzi [Men11] or Frikha and Li [FL17] for some extensions of this technique in other directions. The key point in the mentioned papers consists in performing the first step of a perturbation method for Markov semigroups, known as the parametrix technique, such as exposed in McKean and Singer [MS67], see also Konakov and Mammen [KM00], for the expansion in infinite series of a transition density. In order to do this, one first has to approximate the original system by a simple process for which the well-posedness holds and that does not depend on the original process. This simple approximation process is usually obtained by removing the drift and freezing the diffusion coefficient (with respect to the space variable) in the original dynamics so that its transition density as well as its derivatives can be explicitly estimated. Finally, one crucially has to benefit from the smoothing property of the underlying parametrix kernel, which reflects the quality of the approximation procedure of the original dynamics. However, a direct application of this argument to the McKean-Vlasov SDE (1.1) does not work since the approximation process, obtained by removing the drift and freezing the diffusion coefficient still depends on the original dynamics precisely through the non-linearity induced by the law. Let us also note that a strategy consisting in freezing the measure argument in the dynamics of the original process seems quite unclear and unreasonable since no noise is added in the measure direction and so there is no hope to achieve a smoothing property.

Our strategy to tackle the well-posedness of the martingale problem consists in enlarging the space on which the perturbation argument is performed in order to take into account the non-linearity. More precisely, the underlying space in our analysis is $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. In this new setting, we will need a chain rule formula for a map defined on the Wasserstein space along a flow of probability measures. The chain rule formula that we employ here is the one established by Chassagneux & al. [CCD14], see also Carmona and Delarue [CD18], once we prove the well-posedness and the adequate smoothness of the density of the approximation process. In comparison with the aforementioned results, our approach allows us to deal with non linear diffusion coefficient satisfying a mild structural assumption[1] as well as very mild regularity hypotheses with respect to the space and measure variables, which, to the best of our knowledge, appears to be new. Also, the smoothness assumption on the coefficient b can be weaken to include bounded measurable (with some continuity with respect to the measure variable) drift using some approximation argument but we do not pursue this goal here. By adding a Lispchitz continuity assumption in space on the diffusion coefficient, we derive through usual strong uniqueness results on linear SDE the well-posedness in the strong sense of the SDE (1.1).

The well-posedness of the martingale problem then allows us to investigate in turn the regularity properties of the transition density associated to equation (1.1) and to establish some Gaussian estimates for its derivatives. Some partial results related to the smoothing properties of McKean-Vlasov SDEs have been obtained by Chaudru de Raynal [CdR15], Baños [Bañ18], Crisan and McMurray [CM17]. In [CdR15], such type of bound, under same kind of smoothness assumptions on the coefficients, have been obtained in a regularized framework for McKean-Vlasov SDE (uniformly on the regularization procedure) with scalar interaction only. In [Bañ18], a Bismut-Elworthy-Li formula is proved for a similar equation under the assumption that both the drift and the diffusion matrix are continuously differentiable with bounded Lipschitz derivatives in both variables and the diffusion matrix is uniformly elliptic. In [CM17], using Malliavin calculus techniques, the authors proved several integration by parts formulae for the decoupled dynamics associated to the equation (1.1) from which stem several estimates on the

[1]Let us emphasize that, in this case, only the diffusion coefficient has to satisfy this structural assumption and not the drift.

associated density and its derivatives under smoothness of the coefficients b, σ in the uniform elliptic setting and when the initial law in (1.1) is a Dirac mass.

Here, we will investigate the smoothness properties of the density of both random variables X_t^{ξ} and $X_t^{x,[\xi]}$ (given by the unique associated decoupled flow once the well-posedness for (1.1) has been established) under mild regularity assumptions on the coefficients, namely b and $a = \sigma\sigma^*$ are assumed to be continuous, bounded, Hölder continuous in space and a is uniformly elliptic. In this case, both the drift and diffusion coefficients have to satisfy some structural assumption with respect to the measure variable which guarantees that one can benefit from the smoothing property of the underlying density itself, making the analysis more stringent than in the standard linear setting Friedman [Fri64], [Fri11]. We nevertheless emphasize that even under that structural assumption our result includes coefficients with less than Lipschitz regularity w.r.t. the Wasserstein distance. We eventually establish some Gaussian type estimates for both densities and their derivatives.

Finally, the previous smoothing properties of the densities enable us to investigate classical solutions for a class of linear parabolic PDEs on the Wasserstein space, namely

$$\begin{cases} (\partial_t + \mathcal{L}_t)U(t,x,\mu) = f(t,x,\mu) & \text{for } (t,x,\mu) \in [0,T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d), \\ U(T,x,\mu) = h(x,\mu) & \text{for } (x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d), \end{cases} \tag{1.2}$$

where the source term $f : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and the terminal condition $h : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ are some given functions and the operator $\mathcal{L}_t$ is defined by

$$\begin{aligned} \mathcal{L}_t g(x,\mu) &= \sum_{i=1}^d b_i(t,x,\mu)\partial_{x_i} g(x,\mu) + \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,x,\mu)\partial_{x_i,x_j} g(x,\mu) \\ &\quad + \int \left\{ \sum_{i=1}^d b_i(t,z,\mu)[\partial_\mu g(x,\mu)(z)]_i + \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,z,\mu)\partial_{z_i}[\partial_\mu g(x,\mu)(z)]_j \right\} \mu(dz) \end{aligned} \tag{1.3}$$

and acts on sufficiently smooth test functions $g : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and $a = \sigma\sigma^*$ is uniformly elliptic. Though the first part of the operator appearing in the right-hand side of (1.3) is quite standard, the second part is new and involves the derivative of the test function with respect to the measure variable μ in the sense introduced by P.-L. Lions in his seminal lectures at *the Collège de France*, see [Lio14]. We briefly present this notion of differentiation on the Wasserstein space in Section 2.1 together with the chain rule formula established in [CCD14], see also Carmona and Delarue [CD18], for the flow of measures generated by the law of an Itô process. Classical solutions for PDEs of the form (1.2) have already been investigated in the literature using different methods and under various settings Buckdhan et al. [BLPR17] (for $f \equiv 0$), [CCD14] and very recently [CM17] (for $f \equiv 0$). We also refer the reader to the pedagogical paper Bensoussan et al. [BFY17] for a discussion of the different point of views in order to derive PDEs on the Wasserstein space and their applications.

In the classical diffusion setting, provided the coefficients b and σ and the terminal condition h are smooth enough (with bounded derivatives), it is now well-known that the solution to the related linear Kolmogorov PDE is smooth (see e. g. Krylov [Kry99]). In [BLPR17], the authors proved a similar result in the case of the linear PDE (1.2) (with $f \equiv 0$) and Chassagneux et al. [CCD14] reached the same conclusion for a non-linear version also known as the Master equation. In this sense, the solution of the considered PDE preserves the regularity of the terminal condition. Still in the standard diffusion setting, it is known that one can weaken the regularity assumption on h if one can benefit from the smoothness of the underlying transition density. Indeed in this case, $u(t,x) = \int h(y)\, p(t,T,x,y)\, dy$, $y \mapsto p(t,T,x,y)$ being the density of the (standard) SDE taken at time T and starting from x at time t. However, in order to benefit from this regularizing property, one has to assume that the associated operator $\mathcal{L}$ satisfies some non-degeneracy assumption. When the coefficients b, $a = \sigma\sigma^*$ are bounded measurable and Hölder continuous in space (uniformly in time) and if a is unformly elliptic, it is known (see e.g. [Fri64]) that the linear Kolmogorov PDE admits a fundamental solution so that the unique classical solution exists when the terminal condition h is not differentiable but only continuous. In the seminal paper [Hör67], Hörmander gave a sufficient condition for a second order linear Kolmogorov PDE with smooth coefficients to be hypoelliptic. Thus, if Hörmander's condition is satisfied then the unique classical solution exists even if the terminal condition is not smooth. Note that this condition is known to be nearly necessary since in the non-hypoelliptic regime, even in the case of smooth coefficients, there exists counterexample to the regularity preservation of the terminal condition, see e.g. Hairer and al.

[HHJ15].

The recent paper [CM17] provides the first result in this direction for the PDE (1.2) without source term and for not differentiable terminal condition h using Malliavin calculus techniques under the assumption that the time-homogeneous coefficients b, σ are smooth with respect to the space and measure variables. In particular, the function h has to belong to a certain class of (possibly non-smooth) functions for which Malliavin integration by parts can be applied in order to retrieve the differentiability of the solution in the measure direction. This kind of condition appears to be natural since one cannot expect the solution of the PDE (1.2) to preserve regularity in the measure variable as it is the case in the space argument, see Example 5.1 in [CM17] for more details on this loss of regularity. Imposing our structural assumption on the data f and h and on the coefficients b and a, we derive a theory on the existence and uniqueness of classical solutions for the PDE (1.2) which is analogous to the one considered in Chapter 1 [Fri64] for linear parabolic PDEs. The drift and diffusion coefficients b and a are assumed to be bounded, Hölder continuous in space and they both satisfy together with the terminal condition h and the source term f a structural assumption with respect to the measure argument. The central idea behind the latter assumption, which will be pursued throughout the paper, is to be able to take advantage of the smoothing effect of the underlying noise and thus to weaken the regularity assumptions on the coefficients, especially with respect to the measure argument. This idea could seems strange at first sight since there is no Brownian motion and no Laplacian acting on the measure variable in (1.1) and (1.2). Moreover, as previously mentioned, there is no hope to take advantage of the smoothing effect in that direction in full generality. This is the reason why we are led to consider a specific class of law dependence, allowing to recover the spatial smoothing effect of the noise (or equivalently of the Laplacian) in the measure direction. Fortunately, it appears that such a structural assumption is not that restrictive in practice since it holds for a very large class of interactions considered so far in the literature such as multiple scalar interactions, multiple order interactions, polynomials on Wasserstein space.

The paper is organized as follows. The basic notion of differentiation on the Wasserstein space with an emphasis on the chain rule and on the structural class of maps that will play a central role in our analysis are presented in Section 2. The general set-up together with the assumptions and the main results are described in Section 3. The well-posedness of the martingale problem associated to the SDE (1.1) is tackled in Section 4. The existence and the smoothness properties of its transition density are investigated in Section 5. Finally, classical solutions to the Cauchy problem related to the PDE (1.2) are studied in Section 6. The proof of some useful technical results are given in Appendix.

Notations: In the following we will denote by C and K some generic positive constants that may depend on the coefficients b and σ. We reserve the notation c for constants depending on $|\sigma|_\infty$ and λ (see assumption (**HE**) in Section 3) but not on the time horizon T. Moreover, the value of both C, K or c may eventually change from line to line.

We will denote by $\mathcal{P}(\mathbb{R}^d)$ the space of probability measures on $\mathbb{R}^d$ and by $\mathcal{P}_2(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ the space of probability measures with finite second moment.

For a positive variance-covariance matrix Σ, the function $y \mapsto g(\Sigma, y)$ stands for the d-dimensional Gaussian kernel with Σ as covariance matrix $g(\Sigma, x) = (2\pi)^{-\frac{d}{2}}(\det \Sigma)^{-\frac{1}{2}} \exp(-\frac{1}{2}\langle \Sigma^{-1}x, x\rangle)$. We also define the first and second order Hermite polynomials: $H_1^i(\Sigma, x) := -(\Sigma^{-1}x)_i$ and $H_2^{i,j}(\Sigma, x) := (\Sigma^{-1}x)_i(\Sigma^{-1}x)_j - (\Sigma^{-1})_{i,j}$, $1 \le i, j \le d$ which are related to the previous Gaussian density as follows $\partial_{x_i} g(\Sigma, x) = H_1^i(\Sigma, x)g(\Sigma, x)$, $\partial^2_{x_i,x_j} g(\Sigma, x) = H_2^{i,j}(\Sigma, x)g(\Sigma, x)$. Also, when $\Sigma = cI_d$, for some positive constant c, the latter notation is simplified to $g(c, x) := (1/(2\pi c))^{d/2} \exp(-|x|^2/(2c))$.

One of the key inequality that will be used intensively in this work is the following: for any $p, q > 0$ and $x \in \mathbb{R}$, $|x|^p e^{-qx^2} \le (p/(2qe))^{p/2}$. As a direct consequence, we obtain the *space-time inequality*,

$$\forall p,\, c > 0, \quad |x|^p g(ct, x) \le C t^{p/2} g(c't, x) \tag{1.4}$$

which in turn gives the *standard Gaussian estimates* for the first and second order derivatives of Gaussian density, namely

$$\forall c > 0,\, |H_1^i(ct, x)| \le \frac{C}{t^{\frac{1}{2}}} g(c't, x) \quad \text{and} \quad |H_2^{i,j}(ct, x)| \le \frac{C}{t} g(c't, x) \tag{1.5}$$

for some positive constants C, c'. Since we will employ it quite frequently, we will often omit to mention it explicitly at some places. We finally define the Mittag-Leffler function $E_{\alpha,\beta}(z) := \sum_{n\ge 0} z^n/\Gamma(\alpha n + \beta)$, $z \in \mathbb{R}$, $\alpha,\ \beta > 0$.

2. Preliminaries: Differentiation on the Wasserstein space and structural class

2.1. **Differentiation on the Wasserstein space.** In this section, we present the reader with a brief overview of the regularity notion used when working with mappings defined on $\mathcal{P}_2(\mathbb{R}^d)$. We refer the reader to Lions' seminal lectures [Lio14], to Cardaliaguet's lectures notes [Car13] or to Chapter 5 of Carmona and Delarue's monograph [CD18] for a more complete and detailed exposition. The space $\mathcal{P}_2(\mathbb{R}^d)$ is equipped with the 2-Wasserstein metric

$$W_2(\mu,\nu) = \inf_{\pi\in\mathcal{P}(\mu,\nu)} \left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \, \pi(dx,dy) \right)^{\frac{1}{2}}$$

where, for given $\mu,\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mathcal{P}(\mu,\nu)$ denotes the set of measures on $\mathbb{R}^d\times\mathbb{R}^d$ with marginals μ and ν.

The strategy of Lions consists in considering the canonical lift of the real-valued function $U : \mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto U(\mu)$ into a function $\mathcal{U} : \mathbb{L}_2 \ni Z \mapsto \mathcal{U}(Z) = U([Z]) \in \mathbb{R}$, $(\Omega,\mathcal{F},\mathbb{P})$ standing for an atomless probability space, with Ω a Polish space, $\mathcal{F}$ its Borel σ-algebra, $\mathbb{L}_2 := \mathbb{L}_2(\Omega,\mathcal{F},\mathbb{P},\mathbb{R}^d)$ standing for the space of $\mathbb{R}^d$-valued random variables defined on Ω with finite second moment and Z being a random variable with law μ. Taking advantage of the Hilbert structure of the $\mathbb{L}_2$ space, the function U is then said to be differentiable at $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ if its canonical lift $\mathcal{U}$ is Fréchet differentiable at some point Z such that $[Z]=\mu$. In that case, its gradient is denoted by $D\mathcal{U}$. Thanks to Riezs' representation Theorem, we can identify $D\mathcal{U}$ as an element of $\mathbb{L}^2$. It then turns out that $D\mathcal{U}$ is a random variable which is $\sigma(Z)$-measurable and given by a function $DU(\mu)(.)$ from $\mathbb{R}^d$ to $\mathbb{R}^d$, which depends on the law μ of Z and satisfying $DU(\mu)(.)\in\mathbb{L}^2(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\mu;\mathbb{R}^d)$. Since we will work with mappings U depending on several variables, we will adopt the notation $\partial_\mu U(\mu)(.)$ in order to emphasize that we are taking the derivative of the map U with respect to its measure argument. Thus, inspired by [CD18], the L-derivative (or L-differential) of U at μ is the map $\partial_\mu U(\mu)(.):\mathbb{R}^d\ni v\mapsto \partial_\mu U(\mu)(v)\in\mathbb{R}^d$, satisfying $D\mathcal{U}=\partial_\mu U(\mu)(Z)$.

It is important to note that this representation holds irrespectively of the choice of the original probability space $(\Omega,\mathcal{F},\mathbb{P})$. In what follows, we will only consider functions which are $\mathcal{C}^1$, that is, functions for which the associated canonical lift is $\mathcal{C}^1$ on $\mathbb{L}^2$. We will also restrict our consideration to the class of functions which are $\mathcal{C}^1$ and for which there exists a continuous version of the mapping $\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(\mu,v)\mapsto\partial_\mu U(\mu)(v)\in\mathbb{R}^d$. It then appears that this version is unique. We straightforwardly extend the above discussion to $\mathbb{R}^d$-valued or $\mathbb{R}^d\otimes\mathbb{R}^d$-valued maps U defined on $\mathcal{P}_2(\mathbb{R}^d)$, component by component.

In order to perform the perturbation argument on $\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ and to tackle the PDE (1.2) on the Wasserstein space, we need a chain rule formula for $(U(t,Y_t,[X_t]))_{t\geq0}$, where $(X_t)_{t\geq0}$ and $(Y_t)_{t\geq0}$ are two Itô processes defined for sake of simplicity on the same probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ assumed to be equipped with a right-continuous and complete filtration $\mathbb{F}=(\mathcal{F}_t)_{t\geq0}$. Their dynamics are given by

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s,\ X_0\in\mathbb{L}_2, \tag{2.1}$$

$$Y_t = Y_0 + \int_0^t \eta_s\,ds + \int_0^t \gamma_s\,dW_s \tag{2.2}$$

where $W=(W_t)_{t\geq0}$ is an $\mathbb{F}$-adapted d-dimensional Brownian, $(b_t)_{t\geq0}$, $(\eta_t)_{t\geq0}$, $(\sigma_t)_{t\geq0}$ and $(\gamma_t)_{t\geq0}$ are $\mathbb{F}$-progressively measurable processes, with values in $\mathbb{R}^d$, $\mathbb{R}^d$, $\mathbb{R}^d\otimes\mathbb{R}^d$ and $\mathbb{R}^{d\times q}$ respectively, satisfying the following conditions

$$\forall T>0,\quad \mathbb{E}\Big[\int_0^T (|b_t|^2+|\sigma_t|^4)\,dt\Big]<\infty \text{ and } \mathbb{P}\left(\int_0^T (|\eta_t|+|\gamma_t|^2)\,dt<+\infty\right)=1. \tag{2.3}$$

We now introduce two classes of functions we will work with throughout the paper.

Definition 2.1. (The space $\mathcal{C}^{p,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$, for $p=0,\,1$) Let $T>0$ and $p\in\{0,1\}$. The continuous function $U:[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ is in $\mathcal{C}^{p,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$ if the following conditions hold:

(i) For any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, the mapping $[0,T]\times\mathbb{R}^d\ni(t,x)\mapsto U(t,x,\mu)$ is in $\mathcal{C}^{p,2}([0,T]\times\mathbb{R}^d)$ and the functions $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(t,x,\mu)\mapsto\partial_t^p U(t,x,\mu)$, $\partial_x U(t,x,\mu)$, $\partial_x^2 U(t,x,\mu)$ are continuous.

(ii) For any $(t,x)\in[0,T]\times\mathbb{R}^d$, the mapping $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto U(t,x,\mu)$ is continuously L-differentiable and for any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, we can find a version of the mapping $\mathbb{R}^d\ni v\mapsto\partial_\mu U(t,x,\mu)(v)$ such that the mapping $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(t,x,\mu,v)\mapsto\partial_\mu U(t,x,\mu)(v)$ is locally bounded and is continuous at any (t,x,μ,v) such that $v\in\mathrm{Supp}(\mu)$.

(iii) For the version of $\partial_\mu U$ mentioned above and for any (t,x,μ) in $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, the mapping $\mathbb{R}^d \ni v \mapsto \partial_\mu U(t,x,\mu)(v)$ is continuously differentiable and its derivative $\partial_v[\partial_\mu U(t,x,\mu)](v) \in \mathbb{R}^{d\times d}$ is jointly continuous in (t,x,μ,v) at any point (t,x,μ,v) such that $v \in \mathrm{Supp}(\mu)$.

Remark 2.2. We will also consider the space $\mathcal{C}^{1,p}([0,T]\times\mathcal{P}_2(\mathbb{R}^d))$ for $p = 1, 2$, where we adequately remove the space variable in the definition 2.1. We will say that $U \in \mathcal{C}^{1,1}([0,T]\times\mathcal{P}_2(\mathbb{R}^d))$ if U is continuous, $t \mapsto U(t,\mu) \in \mathcal{C}^1([0,T])$ for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $(t,\mu) \mapsto \partial_t U(t,\mu)$ being continuous and if for any $t \in [0,T]$, $\mu \mapsto U(t,\mu)$ is continuously L-differentiable such that we can find a version of $v \mapsto \partial_\mu U(t,\mu)(v)$ satisfying: $(t,\mu,v) \mapsto \partial_\mu U(t,\mu)(v)$ is locally bounded and continuous at any (t,μ,v) satisfying $v \in \mathrm{Supp}(\mu)$.

We will say that $U \in \mathcal{C}^{1,2}([0,T]\times\mathcal{P}_2(\mathbb{R}^d))$ if $U \in \mathcal{C}^{1,1}([0,T]\times\mathcal{P}_2(\mathbb{R}^d))$ and for the version of $\partial_\mu U$ previously considered, for any $(t,\mu) \in [0,T]\times\mathcal{P}_2(\mathbb{R}^d)$, the mapping $\mathbb{R}^d \ni v \mapsto \partial_\mu U(t,\mu)(v)$ is continuously differentiable and its derivative $\partial_v[\partial_\mu U(t,\mu)](v) \in \mathbb{R}^{d\times d}$ is jointly continuous in (t,μ,v) at any point (t,μ,v) such that $v \in \mathrm{Supp}(\mu)$.

With the above definitions, we can now provide the chain rule formula on the Wasserstein space that will be used intensively in our analysis.

Proposition 2.1 ([CD18], Proposition 5.102)**.** Let X and Y be two Itô processes, with respective dynamics (2.1) and (2.2), satisfying (2.3). Assume that $U \in \mathcal{C}^{1,2,2}([0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$ in the sense of Definition 2.1 such that for any compact set $\mathcal{K} \subset \mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$,

$$\sup_{(t,x,\mu)\in[0,T]\times\mathcal{K}} \left\{ \int_{\mathbb{R}^d} |\partial_\mu U(t,x,\mu)(v)|^2\,\mu(dv) + \int_{\mathbb{R}^d} |\partial_v[\partial_\mu U(t,x,\mu)](v)|^2\,\mu(dv) \right\} < \infty. \tag{2.4}$$

Then, $\mathbb{P}$-a.s., $\forall t \in [0,T]$, one has

$$\begin{aligned} U(t,Y_t,[X_t]) &= U(0,Y_0,[X_0]) + \int_0^t \partial_x U(s,Y_s,[X_s])\,.\gamma_s\,dW_s \\ &+ \int_0^t \left\{ \partial_s U(s,Y_s,[X_s]) + \partial_x U(s,Y_s,[X_s]).\eta_s + \frac{1}{2} Tr(\partial_x^2 U(s,Y_s,[X_s])\gamma_s\gamma_s^T) \right\} ds \\ &+ \int_0^t \left\{ \widetilde{\mathbb{E}}\big[\partial_\mu U(s,Y_s,[X_s])(\widetilde{X}_s).\widetilde{b}_s\big] + \frac{1}{2}\widetilde{\mathbb{E}}\big[Tr(\partial_v[\partial_\mu U(s,Y_s,[X_s])](\widetilde{X}_s)\widetilde{a}_s)\big] \right\} ds \end{aligned} \tag{2.5}$$

where the Itô process $(\widetilde{X}_t,\widetilde{b}_t,\widetilde{\sigma}_t)_{0\le t\le T}$ is a copy of the original process $(X_t,b_t,\sigma_t)_{0\le t\le T}$ defined on a copy $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$ of the original probability space $(\Omega,\mathcal{F},\mathbb{P})$.

2.2. **Structural class.** In this part, we present the structural class of maps defined on $\mathcal{P}_2(\mathbb{R}^d)$ that will have a key role in our analysis. As we already said, assuming that the coefficients b and a belong to this class will allow us to differentiate the density associated to a McKean-Vlasov SDE with respect to its measure argument. Later on, this will enable us to benefit from the smoothing property of the underlying heat kernel and thus in turn to perform the perturbation argument on $\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ as well as to provide a well-posedness theory for classical solutions to the related PDE (1.2).

In order to foster the understanding of the main idea, let us start with a very simple example of such a function by considering what is called a first order interaction. To be more specific, a function $h : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ satisfies a first order interaction if it is of following form

$$h(\mu) = \int \widetilde{h}(z)\mu(dz).$$

for some measurable function $\widetilde{h} : \mathbb{R}^d \to \mathbb{R}$. Recalling the definition of L-derivative, one immediately sees that $\partial_\mu h(\mu)(v) = \partial_x \widetilde{h}(v)$ so that the function h is, in full generality, differentiable if and only if $\widetilde{h}$ is differentiable. It is hence clear that, if one considers the SDE (1.1) with $d = q = 1$, $b \equiv 0$ and $\sigma \equiv 0$, one cannot expect the map $\mu \mapsto h([X_T^{t,\xi}])$ to be differentiable without assuming that $\widetilde{h}$ is smooth.

Assuming now that $\sigma > 0$, let us show how the spatial random perturbation of $X_T^{t,\xi} = \xi + \sigma(W_T - W_t)$ by the Brownian motion (which is assumed to be independent of ξ) allows to regularize such a map h through $\widetilde{h}$. Let us assume for sake of simplicity that $\widetilde{h}$ is bounded. Setting for simplicity $\Theta(t,\mu)(dz) := [X_T^{t,\xi}](dz) = \int g(a(T-t), z-x)\,\mu(dx)\,dz$, $a = \sigma^2$, by Fubini's theorem, we have that $h(\Theta(t,\mu)) = \int \widetilde{h}(z)\,\Theta(t,\mu)(dz) = \int\int \widetilde{h}(z)\,g(a(T-t),z-x)\,\mu(dx)\,dz = \int\int \widetilde{h}(z)\,g(a(T-t),z-x)\,dz\,\mu(dx)$ so that

$\mu \mapsto h(\Theta(t,\mu))$ is now smooth with derivatives:

$$\begin{aligned}\partial_\mu h(\Theta(t,\mu))(v) &= \partial_\mu \left[\int\int \widetilde{h}(z)\, g(a(T-t), z-x)\, dz\, \mu(dx)\right](v)\\ &= \partial_v \left[\int \widetilde{h}(z) g(a(T-t), z-v)\, dz\right]\\ &= \int \widetilde{h}(z)\, (-H_1.g)(a(T-t), z-v)\, dz.\end{aligned}$$

The crucial point here is that the map $\Theta : [0,T] \times \mathcal{P}_2(\mathbb{R}^d) \ni (t,\mu) \mapsto \Theta(t,\mu) \in \mathcal{P}_2(\mathbb{R}^d)$ is smooth and that the linearity with respect to the variable μ of h allows to transfer the L-derivative on h to $\Theta(t,\mu)$ which in turn reads as a space derivative on the underlying Gaussian kernel. Note that the map $[0,T) \ni t \mapsto h(\Theta(t,\mu))$ is also differentiable with

$$\begin{aligned}\partial_t h(\Theta(t,\mu)) &= \partial_s \Big[\int\int \widetilde{h}(z)\, g(a(T-s), z-x)\, dz\, \mu(dx)\Big]_{s=t}\\ &= -a \int\int \widetilde{h}(z)\, \partial_1 g(a(T-t), z-x)\, dz\, \mu(dx)\\ &= -\frac{a}{2} \int\int \widetilde{h}(z)\, (H_2.g)(a(T-t), z-x)\, dz\, \mu(dx).\end{aligned}$$

Such a property naturally leads to the following general definition for our class of coefficients.

Definition 2.3. [Class of functions $(\mathbf{CS})$ and $(\mathbf{CS}_+)$] A map $h : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is in the class $(\mathbf{CS})$ if for any $T > 0$, for any mapping $(t,x,\mu) \mapsto p(\mu,t,T,x,z) \in \mathcal{C}^{1,2,2}([0,T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$, $z \mapsto p(\mu,t,T,x,z)$ being a density function, such that $(p(\mu,t,T,.,z)\sharp\mu)dz \in \mathcal{P}_2(\mathbb{R}^d)$ and for any compact set $\mathcal{K} \subset [0,T) \times \mathcal{P}_2(\mathbb{R}^d) \times (\mathbb{R}^d)^2$, for any $n = 0, 1$,

$$\int \sup_{(t,\mu,x,v)\in\mathcal{K}} \big\{|\partial_t^n p(\mu,t,T,x,z)| + |\,\partial_x^{1+n} p(\mu,t,T,x,z)| + |\partial_v^n[\partial_\mu p(\mu,t,T,x,z)](v)|\big\}\, dz < \infty, \tag{2.6}$$

denoting by $\Theta(t,\mu) : [0,T) \times \mathcal{P}_2(\mathbb{R}^d) \ni (t,\mu) \mapsto \Theta(t,\mu)(dz) \in \mathcal{P}_2(\mathbb{R}^d)$ with $\Theta(t,\mu)(dz) = (p(\mu,t,T,.,z)\sharp\mu)(dz) = \int_{\mathbb{R}^d} p(\mu,t,T,x,z)\mu(dx)\, dz$, one has:

- the map $[0,T) \times \mathcal{P}_2(\mathbb{R}^d) \ni (t,\mu) \mapsto h(\Theta(t,\mu)) \in \mathcal{C}^{1,2}([0,T) \times \mathcal{P}_2(\mathbb{R}^d))$,
- the Lions and time derivatives satisfy for $n = 0, 1$:

$$\begin{aligned}\partial_v^n[\partial_\mu h(\Theta(t,\mu))](v) &= \partial_v^n \Big[\partial_\nu \Big[\int \widetilde{h}(z, \Theta(t,\mu))\, \Theta(t,\nu)(dz)\Big]_{|\nu=\mu}\Big](v)\\ &= \partial_v^n \Big[\partial_\nu \Big[\int\int \widetilde{h}(z, \Theta(t,\mu))\, p(\nu,t,T,x,z)\, dz\, \nu(dx)\Big]_{|\nu=\mu}\Big](v) \qquad (2.7)\\ \partial_t h(\Theta(t,\mu)) &= \partial_s \Big[\int \widetilde{h}(z, \Theta(t,\mu))\, \Theta(s,\nu)(dz)\Big]_{|s=t}\\ &= \partial_s \Big[\int \widetilde{h}(z, \Theta(t,\mu))\, p(\mu,s,T,x,z)\, dz\, \mu(dx)\Big]_{|s=t}. \qquad (2.8)\end{aligned}$$

  for some bounded continuous function $\widetilde{h} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$.

We will say that the map h is in the class $(\mathbf{CS}_+)$ if h is in $(\mathbf{CS})$ and the two following conditions $(\mathbf{CS}_+)_1$ and $(\mathbf{CS}_+)_2$ are satisfied:

$(\mathbf{CS}_+)_1$ The map $z \mapsto \widetilde{h}(z,\mu)$ is η-Hölder continuous, for some $\eta \in (0,1]$, with modulus denoted by $[\widetilde{h}]_H$, uniformly with respect to μ.

$(\mathbf{CS}_+)_2$ There exists a map $\widetilde{H} : (\mathbb{R}^d)^2 \times (\mathcal{P}_2(\mathbb{R}^d))^2 \to \mathbb{R}^N$, for some positive integer N, such that $z' \mapsto \widetilde{H}(z,z',\mu,\nu)$ is η-Hölder continuous, with modulus denoted by $[\widetilde{H}]_H$, uniformly with respect to the other variables and satisfying:

$$\forall (z,\mu,\nu) \in \mathbb{R}^d \times (\mathcal{P}_2(\mathbb{R}^d))^2, \quad |\widetilde{h}(z,\nu) - \widetilde{h}(z,\mu)| \le \Big|\int \widetilde{H}(z,z',\nu,\mu)\, (\nu-\mu)(dz')\Big|. \tag{2.9}$$

Remark 2.4. ∘ Importantly, we note that the above definition does not impose any smoothness assumption on the function h but rather *a structure* on the derivative (when it exists) along a smooth flow $(t,\mu) \mapsto \Theta(t,\mu)$ of probability measures of $\mathcal{P}_2(\mathbb{R}^d)$. As already said, in what follows, the map Θ will be the one generated by the unique weak solution of an SDE of the form (1.1), that is, one is interested in the smoothness of $[0,T) \times \mathcal{P}_2(\mathbb{R}^d) \ni (t,\mu) \mapsto h([X_T^{t,\xi}])$ so that we will often consider $\Theta : [0,T) \times \mathcal{P}_2(\mathbb{R}^d) \ni (t,\mu) \mapsto [X_T^{t,\xi}] \in \mathcal{P}_2(\mathbb{R}^d)$.

∘ For functions $h : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $h : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$, we will straightforwardly extend the above definition to each component and still denote $\widetilde{h} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\widetilde{h} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ the corresponding maps.

∘ Finally, when a function h belonging to the class $(\mathbf{CS})$ or $(\mathbf{CS}_+)$ does depend on other variables, it is implicitly assumed that the continuity as well as the Hölder regularity property stated in Definition 2.3 hold uniformly. For example, if one assumes that $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto a(t,x,\mu) \in \mathbb{R}^{d\times d}$ is in the class $(\mathbf{CS}_+)$ then the corresponding map $z \mapsto \widetilde{a}(t,x,z,\mu)$ appearing in conditions (2.7) and (2.8) is assumed to be η-Hölder, with modulus denoted by $[\widetilde{a}]_H$, uniformly with respect to $(t,x,\mu) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. The same remark holds for the map $z' \mapsto \widetilde{A}(t,x,z,z',\nu,\mu)$ appearing in the condition (2.9) and we will denote by $[\widetilde{A}]_H$ its Hölder modulus. In what follows, we will also denote by $\widetilde{b}$ and $\widetilde{B}$ (and by $[\widetilde{b}]_H$, $[\widetilde{B}]_H$ their Hölder modulus) the two maps appearing respectively in the conditions (2.7), (2.8) and (2.9) as soon as one assumes that $\mu \mapsto b(t,x,\mu)$ is in $(\mathbf{CS})$ or $(\mathbf{CS}_+)$.

Typical examples of functions in the classes (CS) and ($\mathbf{CS}_+$).

Let us illustrate this definition by giving some explicits examples of functions belonging to the classes $(\mathbf{CS})$ and $(\mathbf{CS}_+)$. In the following, h denotes a map from $\mathcal{P}_2(\mathbb{R}^d)$ to $\mathbb{R}$. We can straightforwardly consider their multidimensional version by applying the above remark.

(1) First order interaction belongs to $(\mathbf{CS})$. We say that h satisfies a first order interaction if it is of following form: for some bounded continuous function $\bar{h} : \mathbb{R}^d \to \mathbb{R}$, one has
$$h(\mu) = \int \bar{h}(y)\mu(dy).$$

(2) N order interaction belongs to $(\mathbf{CS})$. We say that h satisfies an N order interaction if it is of following form: for some bounded continuous $\bar{h} : \mathbb{R}^N \to \mathbb{R}$, one has
$$h(\mu) = \int \cdots \int \bar{h}(y_1, \cdots, y_N)\, \mu(dy_1) \cdots \mu(dy_N).$$

(3) Polynomials on the Wasserstein space belong to $(\mathbf{CS})$. We say that a function f is a polynomial on the Wasserstein space if there exist some real-valued bounded continuous functions $\bar{h}_1, \cdots, \bar{h}_N$ defined on $\mathbb{R}^d$ such that
$$h(\mu) = \prod_{i=1}^{N} \int \bar{h}_i(z)\mu(dz).$$

(4) Scalar interaction belongs to $(\mathbf{CS})$. We say that a function h satisfies a scalar interaction if there exist a continuously differentiable real-valued function $\bar{h}$ defined on $\mathbb{R}^N$ as well as some real-valued bounded continuous functions $\bar{h}_1, \cdots, \bar{h}_N$ defined on $\mathbb{R}^d$ such that
$$h(\mu) = \bar{h}\left(\int \bar{h}_1(y)\,\mu(dy), \cdots, \int \bar{h}_N(y)\,\mu(dy)\right).$$

(5) Sum, product and more generally any smooth composition of N order interactions, polynomials on Wasserstein space or scalar interaction belong to $(\mathbf{CS})$.

Any N order interactions as described in $(1)-(2)$ with bounded $\eta-$Hölder continuous functions $\bar{h}$; any polynomial in Wasserstein space as described in (3) with bounded $\eta-$Hölder continuous $\bar{h}_i$, $i = 1, \cdots, N$ as well as any sum, product or smooth function of these two classes under the aforementioned assumptions belong to the class $(\mathbf{CS}_+)$. For the scalar interaction case, if $\bar{h}$ has a Lipschitz-continuous derivative and if each $\bar{h}_i$, $i = 1, \cdots, N$, are bounded η-Hölder continuous functions, then the function f belongs to $(\mathbf{CS}_+)$.

For the sake of completeness, let us prove the above statement in the case of first order interaction. Similar arguments can be employed for the other aforementioned examples and we therefore leave to the reader the task of writing the remaining technical details for each one of them.

We thus consider a map $(t,x,\mu) \mapsto p(\mu,t,T,x,z) \in \mathcal{C}^{1,2,2}([0,T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ satisfying the conditions of Definition 2.3 for any fixed $T > 0$ and denote $\Theta(t,\mu)(dz) := \int p(\mu,t,T,x,z)\,\mu(dx)\,dz$. By Fubini's theorem, $h(\Theta(t,\mu)) = \int\int \bar{h}(z)\,p(\mu,t,T,x,z)\,dz\mu(dx)$. We start by proving that $(t,\mu) \mapsto h(\Theta(t,\mu))$ is continuous on $[0,T) \times \mathcal{P}_2(\mathbb{R}^d)$. Let $(t_n,\mu_n)_{n\geq 1}$ be a sequence of $[0,T) \times \mathcal{P}_2(\mathbb{R}^d)$ satisfying $\lim_n |t_n - t| = \lim_n W_2(\mu_n,\mu) = 0$. We decompose the difference $\int \bar{h}(z)p(\mu_n,t_n,T,x,z)\,dz\,\mu_n(dx) - \int \bar{h}(z)p(\mu,t,T,x,z)\,dz\,\mu(dx)$ as the sum of two terms, namely

$$\mathrm{I}_n := \int\int \bar{h}(z)p(\mu_n,t_n,T,x,z)\,dz(\mu_n-\mu)(dz), \quad \mathrm{II}_n := \int\int \bar{h}(z)(p(\mu_n,t_n,T,x,z)-p(\mu,t,T,x,z))\,dz\mu(dx)$$

and prove that each term goes to zero as $n \uparrow \infty$. Let us note that from condition (2.6) and Lebesgue's dominated convergence theorem, one directly gets $\lim_n \mathrm{II}_n = 0$. In order to prove that $\lim_n \mathrm{I}_n = 0$, we decompose I_n as the sum of two terms namely

$$\mathrm{I}_n^1 := \int\int \bar{h}(z)p(\mu_n, t_n, T, x, z)\, dz\, \eta_R(x)(\mu_n - \mu)(dx),\ \mathrm{I}_n^2 := \int\int \bar{h}(z)p(\mu_n, t_n, T, x, z)\, dz\, (1-\eta_R)(x)(\mu_n - \mu)(dx)$$

where η_R is a non-negative smooth cutoff function such that $\eta_R(x) = 1$ for $|x| \leq R$ and $\eta_R(x) = 0$ for $|x| \geq 2R$, with $R > 0$. The uniform continuity of the map $\mathcal{K} \times B_{2R} \ni (t, \mu, x) \mapsto \int \bar{h}(z)p(\mu, t, T, x, z)\, dz\, \eta_R(x)$, $\mathcal{K}$ being a compact set of $[0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ and B_{2R} being the closed ball of radius R around the origin, implies that the family of maps $\big\{\int \bar{h}(z)p(\mu, t, T, x, z)\, dz\, \eta_R(x),\ (t, \mu) \in \mathcal{K}\big\}$ is equicontinuous and (2.6) implies its boundedness. By weak convergence of $(\mu_n)_{n\geq 1}$, we thus deduce

$$\lim_n \sup_{(t,\mu')\in\mathcal{K}} \Big|\int\int \bar{h}(z)p(\mu', t, T, x, z)\, dz\, \eta_R(x)\, (\mu_n - \mu)(dx)\Big| = 0$$

so that $\lim_n \mathrm{I}_n^1 = 0$. From the boundedness of $\bar{h}$ and the weak convergence of $(\mu_n)_{n\geq 1}$, we also obtain $\limsup_n |\mathrm{I}_n^2| \leq |\bar{h}|_\infty(\limsup_n \int_{|x|\geq 2R} \mu_n(dx) + \int_{|x|\geq 2R} \mu(dx)) \leq 2|\bar{h}|_\infty \int_{|x|\geq 2R} \mu(dx)$ so that by letting R goes to infinity in the previous inequality we deduce $\lim_n \mathrm{I}_n^2 = 0$. We thus conclude that $(t, \mu) \mapsto h(\Theta(t, \mu))$ is continuous on $[0, T) \times \mathcal{P}_2(\mathbb{R}^d)$.

Moreover, from condition (2.6), by Fubini's and Lebesgue's differentiation theorems, one deduces that $\mu \mapsto \int \bar{h}(z)\, p(\mu, t, T, x, z)\, dz\, \mu(dx)$ is differentiable at any fixed $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ with

$$\begin{aligned}
&\partial_\mu\Big[\int\int \bar{h}(z)p(\mu, t, x, z)\, \mu(dx)\, dz\Big](v)\\
&= \partial_\nu\Big[\int\int \bar{h}(z)p(\nu, t, x, z)\, dz\, \mu(dx)\Big]_{\nu=\mu}(v) + \partial_\nu\Big[\int\int \bar{h}(z)p(\mu, t, T, x, z)\, dz\, \nu(dx)\Big]_{\nu=\mu}(v)\\
&= \int\int \bar{h}(z)\partial_\mu p(\mu, t, T, x, z)(v)\, dz\, \mu(dx) + \int \bar{h}(z)\partial_x p(\mu, t, T, v, z)\, dz.
\end{aligned}$$

In order to prove that each term appearing in the right-hand side of the previous expression is globally continuous w.r.t t, μ, v, one proceeds as previously done by considering a sequence $(t_n, \mu_n, v_n)_{n\geq 1}$ converging to $(t, \mu, v) \in [0, T)\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$. By dominated convergence, one obtains $\lim_n \int \bar{h}(z)\partial_x p(\mu_n, t_n, T, v_n, z)\, dz = \int \bar{h}(z)\partial_x p(\mu, t, T, v, z)\, dz$. Employing the previous decomposition using the cutoff function η_R and the continuity of $(t, x, \mu, v) \mapsto \partial_\mu p(\mu, t, T, x, z)(v)$ as well as condition (2.6), one proves in a completely analogous manner that $\lim \int\int \bar{h}(z)\partial_\mu p(\mu_n, t_n, T, x, z)(v_n)\, dz\, \mu_n(dx) = \int\int \bar{h}(z)\partial_\mu p(\mu, t, T, x, z)(v)\, dz\, \mu(dx)$. Still, from the previous expression, using again Lebesgue's differentiation theorem under condition (2.6), one deduces that the mapping $\mathbb{R}^d \ni v \mapsto \partial_\mu\big[\int\int \bar{h}(z)p(\mu, t, x, z)\, \mu(dx)\, dz\big](v)$ is continuously differentiable with derivative

$$\partial_v\Big[\partial_\mu\Big[\int\int \bar{h}(z)p(\mu, t, x, z)\, \mu(dx)\, dz\Big]\Big](v) = \int\int \bar{h}(z)\partial_v[\partial_\mu p(\mu, t, T, x, z)](v)\, dz\, \mu(dx) + \int \bar{h}(z)\partial_x^2 p(\mu, t, T, v, z)\, dz$$

and the joint continuity of the previous expression in $(t, x, \mu, v) \in [0, T)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$ stems from similar lines of reasonings as those employed before. Finally, one also obtains that $t \mapsto \int\int \bar{h}(z)p(\mu, t, x, z)\, \mu(dx)\, dz$ is continuously differentiable on $[0, T)$ with derivative $\int\int \bar{h}(z)\partial_t p(\mu, t, x, z)\, \mu(dx)\, dz$ being jointly continuous in (t, μ). We thus conclude that $(t, \mu) \mapsto h(\Theta(t, \mu)) \in \mathcal{C}^{1,2}([0, T) \times \mathcal{P}_2(\mathbb{R}^d))$ so that h belongs to $(\mathbf{CS})$ after noting that $\widetilde{h}(z, \mu) = \bar{h}(z)$ for any $(z, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. Once again a similar analysis can be performed for each of the aforementioned examples which demonstrates that both classes $(\mathbf{CS})$ and $(\mathbf{CS}_+)$ cover a large class of non-linear interaction.

To conclude we again emphasize that the main advantage of our class of coefficients is to be able to take advantage of the usual smoothing effect of Gaussian-like kernels. We will intensively exploit this property to establish our regularity results for the density associated to solutions of McKean-Vlasov SDEs. We end this section with a simple result that illustrates this central idea.

Lemma 2.1. Let $X_T^{t,\xi}$ be an $\mathbb{R}^d$-valued random variable with density function $z \mapsto p(\mu, t, T, z) = \int_{\mathbb{R}^d} p(\mu, t, T, x, z)\, \mu(dx)$ such that $[0, T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (t, x, \mu) \mapsto p(\mu, t, T, x, z)$ satisfies the conditions of Definition (2.3).

Then, for any h in (**CS**), the derivatives of the map $(t,\mu)\mapsto h([X_T^{t,\xi}])\in\mathcal{C}^{1,2}([0,T)\times\mathcal{P}_2(\mathbb{R}^d))$ admit the following representations

$$\partial_v^n[\partial_\mu h([X_T^{t,\xi}])] = \int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(v,[X_T^{t,\xi}])]\partial_x^{1+n}p(\mu,t,T,v,z)\,dz + \int\int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(x,[X_T^{t,\xi}])]\partial_v^n[\partial_\mu p(\mu,t,T,x,z)](v)\,dz\,\mu(dx),\ n=0,1, \tag{2.10}$$

$$\partial_t h([X_T^{t,\xi}]) = \int\int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(x,[X_T^{t,\xi}])]\ \partial_t p(\mu,t,T,x,z)\,dz\,\mu(dx). \tag{2.11}$$

Proof. It suffices to notice that from the key relation (2.7), by Fubini's and Lebesgue's differentiation theorems, for any $v_0\in\mathbb{R}^d$, one has

$$\begin{aligned}
\partial_v^n\Big[\partial_\mu h([X_T^{t,\xi}])\Big](v) &= \partial_v^n\Big[\partial_\nu\Big[\int\int\widetilde{h}(z,[X_T^{t,\xi}])p(\nu,t,T,x,z)\,\nu(dx)\,dz\Big]_{|\nu=\mu}\Big](v)\\
&= \partial_v^n\Big[\partial_\nu\Big[\int\int\widetilde{h}(z,[X_T^{t,\xi}])p(\mu,t,T,x,z)\,dz\,\nu(dx)\Big]_{|\nu=\mu}\Big](v)\\
&\quad+\partial_v^n\Big[\partial_\nu\Big[\int\int\widetilde{h}(z,[X_T^{t,\xi}])p(\nu,t,T,x,z)\,dz\,\mu(dx)\Big]_{|\nu=\mu}\Big](v)\\
&= \partial_v^n\Big[\partial_\nu\Big[\int\int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(v_0,[X_T^{t,\xi}])]p(\mu,t,T,x,z)\,dz\,\nu(dx)\Big]_{|\nu=\mu}\Big](v)\\
&\quad+\partial_v^n\Big[\partial_\mu\Big[\int\int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(x,[X_T^{t,\xi}])]p(\nu,t,T,x,z)\,dz\,\mu(dx)\Big]_{|\nu=\mu}\Big](v)\\
&\quad+\partial_v^n\Big[\partial_\nu\Big[\int\int\widetilde{h}(v_0,[X_T^{t,\xi}])p(\mu,t,T,x,z)\,dz\,\nu(dx)\Big]_{|\nu=\mu}\Big](v)\\
&\quad+\partial_v^n\Big[\partial_\nu\Big[\int\int\widetilde{h}(x,[X_T^{t,\xi}])p(\nu,t,T,x,z)\,dz\,\mu(dx)\Big]_{|\nu=\mu}\Big](v)\\
&= \int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(v_0,[X_T^{t,\xi}])]\partial_x^{1+n}p(\mu,t,T,v,z)\,dz\\
&\quad+\int\int[\widetilde{h}(z,[X_T^{t,\xi}])-\widetilde{h}(x,[X_T^{t,\xi}])]\partial_v^n[\partial_\mu p(\mu,t,T,x,z)](v)\,dz\,\mu(dx)
\end{aligned}$$

where we used the fact that the last two terms appearing in the last but one equality are 0 since $z\mapsto p(\mu,s,t,x,z)$ is a density function. The identity (2.10) then follows by taking $v_0=v$ in the previous identity. The relation (2.11) for the time derivative $\partial_t h([X_T^{t,\xi}])$ follows from a similar argument and technical details are omitted.

□

The two representation formulas (2.10) and (2.11) are crucial for the analysis of the smoothness properties of densities associated to McKean-Vlasov SDEs. Indeed, under the additional assumption that h belongs to (**CS**$_+$) and if $(t,\mu,x)\mapsto p(\mu,t,x,z)$ as well as its derivatives satisfy Gaussian-type bounds, they allow thanks to the η-Hölder regularity of $\widetilde{h}$ and the space-time inequality (1.4) to match the diagonal regime of the underlying heat kernel and to benefit from the so-called smoothing property of Gaussian kernels. At this stage, we stop from elaborating on this important idea and postpone the discussion to the appropriate place.

3. Overview, assumptions and main results

3.1. On the well-posedness of the martingale problem related to the SDE (1.1).

We first present the martingale problem associated to equation (1.1).

Definition 3.1. Let $\mu\in\mathcal{P}(\mathbb{R}^d)$. We say that the probability measure $\mathbb{P}$ on the canonical space $\mathcal{C}([0,\infty),\mathbb{R}^d)$ (endowed with the canonical filtration $(\mathcal{F}_t)_{t\geq0}$) with time marginals $(\mathbb{P}(t))_{t\geq0}$, solves the non-linear martingale problem associated to the SDE (1.1) with initial distribution μ at time 0 if the canonical process $(y_t)_{t\geq0}$ satisfies the following two conditions:

(i) $\mathbb{P}(y_0\in\Gamma)=\mu(\Gamma)$, $\Gamma\in\mathcal{B}(\mathbb{R}^d)$.
(ii) For all $f\in\mathcal{C}_b^2(\mathbb{R}^d)$, the process

$$f(y_t)-f(y_0)-\int_0^t\left\{\sum_{i=1}^d b_i(s,y_s,\mathbb{P}(s))\partial_{x_i}f(y_s)+\frac{1}{2}\sum_{i,j=1}^d a_{i,j}(s,y_s,\mathbb{P}(s))\partial^2_{x_i,x_j}f(y_s)\right\}ds \tag{3.1}$$

is a square integrable martingale under $\mathbb{P}$.

Remark 3.2. A similar definition holds by letting the canonical process starts from time t_0 with initial distribution μ, in which case we say that the initial condition is (t_0, μ) and (i) is replaced by the condition: $\mathbb{P}(y(s) \in \Gamma; 0 \leq s \leq t_0) = \mu(\Gamma)$.

Having this definition at hand we now introduce some assumptions on the coefficients:

(**HC**) The drift coefficient $b : [0,\infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ and diffusion coefficient $\sigma : [0,\infty) \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d \otimes \mathbb{R}^q$ are bounded and continuous functions, in the following sense: for all $(i,j) \in \{1, \cdots, d\}$, for all $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$b_i(t,x,\mu) = \lim_{(s,y)\to(t,x)} b_i(s,y,\mu), \quad a_{i,j}(t,x,\mu) = \lim_{(s,y)\to(t,x)} a_{i,j}(s,y,\mu)$$

and, for all $\mu \in \mathcal{P}(\mathbb{R}^d)$, for any T, $R > 0$,

$$\lim_{\pi(\mu,\nu)\downarrow 0} \sup_{0\leq t\leq T, |x|\leq R} |b_i(t,x,\mu) - b_i(t,x,\nu)| = 0,$$
$$\lim_{\pi(\mu,\nu)\downarrow 0} \sup_{0\leq t\leq T, |x|\leq R} |a_{i,j}(t,x,\mu) - a_{i,j}(t,x,\nu)| = 0$$

where π is the Lévy-Prokhorov metric on $\mathcal{P}(\mathbb{R}^d)$.

(**HR**) (i) The function $\mathbb{R}^d \ni x \mapsto a(t,x,\mu) \in \mathbb{R}^d \otimes \mathbb{R}^d$ is uniformly η-Hölder continuous for some $\eta \in (0,1]$, that is,

$$[a]_H := \sup_{t\geq 0,\, x\neq y,\, \mu\in\mathcal{P}_2(\mathbb{R}^d)} \frac{|a(t,x,\mu) - a(t,y,\mu)|}{|x-y|^\eta} < \infty.$$

(ii) There exists a function $A : \mathbb{R}_+ \times (\mathbb{R}^d)^2 \times (\mathcal{P}_2(\mathbb{R}^d))^2 \to \mathbb{R}^N$, for some positive integer N, such that for every $(t,x,\nu,\mu) \in \mathbb{R}_+ \times \mathbb{R}^d \times (\mathcal{P}_2(\mathbb{R}^d))^2$ the function $z \mapsto A(t,x,z,\nu,\mu)$ is η-Hölder continuous, with modulus denoted by $[A]_H$, for some $\eta \in (0,1]$, uniformly with respect to the other variables and satisfying: for all $t \geq 0$, for all $(x,\nu,\mu) \in \mathbb{R}^d \times (\mathcal{P}_2(\mathbb{R}^d))^2$

$$\max_{i,j} |a_{i,j}(t,x,\nu) - a_{i,j}(t,x,\mu)| \leq \Big| \int A(t,x,z,\nu,\mu)(\nu-\mu)(dz) \Big|.$$

(**HE**) The diffusion coefficient is uniformly elliptic, that is, there exists $\lambda \geq 1$ such that for every $(t,\mu) \in [0,\infty) \times \mathcal{P}(\mathbb{R}^d)$ and $(x,\xi) \in (\mathbb{R}^d)^2$, $\lambda^{-1}|\xi|^2 \leq \langle a(t,x,\mu)\xi, \xi\rangle \leq \lambda|\xi|^2$ where $a(t,x,\mu) = (\sigma\sigma^*)(t,x,\mu)$.

Our first main result concerns the well-posedness of the martingale problem associated to the SDE (1.1).

Theorem 3.3. Under (**HC**), (**HR**), (**HE**) and assuming that $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto a(t,x,\mu) \in \mathbb{R}^{d\times d}$ belongs to the class $(\mathbf{CS}_+)$, the martingale problem associated with (1.1) is well-posed for any initial distribution $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. In particular, weak uniqueness in law holds for the SDE (1.1).

When investigating strong well-posedness of non-linear SDE an interesting fact is that, combining uniqueness in law for the non-linear SDE together with strong uniqueness result for the associated linear SDE, *i.e.* the same SDE with time inhomogeneous coefficients, the law argument being now treated as a time-inhomogeneity, immediately yields to strong uniqueness. To be more specific, from the previous well-posedness result we have that any strong solution Y of the SDE (1.1) (if it exists) writes

$$Y_t = \xi + \int_0^t b(s, Y_s, [X_s^\xi])ds + \int_0^t \sigma(s, Y_s, [X_s^\xi])dW_s, \quad [\xi] \in \mathcal{P}_2(\mathbb{R}^d) \tag{3.2}$$

implying that, setting $\widehat{b} : \mathbb{R}^+ \times \mathbb{R}^d \ni (t,y) \mapsto b(t,y,[X_t^\xi]) \in \mathbb{R}^d$ and $\widehat{\sigma} : \mathbb{R}^+ \times \mathbb{R}^d \ni (t,y) \mapsto \widehat{\sigma}(t,y,[X_t^\xi]) \in \mathbb{R}^d \times \mathbb{R}^d$, it solves

$$Y_t = \xi + \int_0^t \widehat{b}(s, Y_s)ds + \int_0^t \widehat{\sigma}(s, Y_s)dW_s, \quad [\xi] \in \mathcal{P}_2(\mathbb{R}^d). \tag{3.3}$$

But this linear SDE is well posed in the strong sense under the additional assumption that the diffusion coefficient $\widehat{\sigma}$ is Lipschitz in space (see [Ver80]). Hence, any strong solutions of (3.2) are equals $\mathbb{P}$-a.s. so that strong well-posedness follows from the Yamada-Watanabe theorem. This gives the following corollary.

Corollary 3.4. *Suppose that assumptions (**HC**), (**HR**), (**HE**) hold and that $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto a(t,x,\mu) \in \mathbb{R}^d \otimes \mathbb{R}^d$ belongs to the class ($\boldsymbol{CS}_+$). Assume moreover that for all (t,μ) in $\mathbb{R}^+ \times \mathcal{P}_2(\mathbb{R}^d)$ the mapping $x \mapsto \sigma(t,x,\mu)$ is Lispchitz continuous uniformly with respect to t and μ. Then, strong uniqueness holds for the SDE* (1.1).

Here are some examples for which our weak and strong uniqueness results apply.

Example 3.5. (First order interaction) We consider the following non-linear SDE with coefficients $b : \mathbb{R}_+ \times (\mathbb{R}^d)^2 \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times (\mathbb{R}^d)^2 \to \mathbb{R}^d \otimes \mathbb{R}^q$:

$$X_t^\xi = \xi + \int_0^t \widetilde{\mathbb{E}}[b(s, X_s^\xi, \widetilde{X}_s^\xi)]ds + \int_0^t \widetilde{\mathbb{E}}[\sigma(s, X_s^\xi, \widetilde{X}_s^\xi)]dW_s, \quad [\xi] \in \mathcal{P}_2(\mathbb{R}^d). \tag{3.4}$$

where the process $(\widetilde{X}_t^\xi)_{t\geq 0}$ is a copy of $(X_t^\xi)_{t\geq 0}$ defined on a copy $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Assume that the functions b and σ are bounded continuous functions; assume that $(x,z) \mapsto \sigma(t,x,z)$ is η-Hölder continuous uniformly with respect to t and that $(\int \sigma(t,x,z)\mu(dz))(\int \sigma(t,x,z)\mu(dz))^*$ is uniformly elliptic, uniformly with respect to the variables t, x, μ.

Then, assumptions of Theorem 3.3 are fulfilled and (3.4) is well posed in the weak sense. If in addition $x \mapsto \sigma(t,x,z)$ is Lipschitz continuous uniformly with respect to t and z then assumptions of Corollary 3.4 are satisfied and strong well posedness holds for (3.4).

An approximation argument that we do not detail here allows to handle the case of a bounded measurable drift function $(t,x,z) \mapsto b(t,x,z)$ for both strong and weak well-posedness.

Example 3.6. (N order interaction) For some positive integer N, we consider the following non-linear SDE with coefficients $b : \mathbb{R}_+ \times (\mathbb{R}^d)^{N+1} \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times (\mathbb{R}^d)^{N+1} \to \mathbb{R}^{d\times q}$:

$$X_t^\xi = \xi + \int_0^t \widetilde{\mathbb{E}}[b(s, X_s^\xi, X_s^{\xi,(1)}, \cdots, X_s^{\xi,(N)})]ds + \int_0^t \widetilde{\mathbb{E}}[\sigma(s, X_s^\xi, X_s^{\xi,(1)}, \cdots, X_s^{\xi,(N)})]dW_s, \tag{3.5}$$

with $[\xi] \in \mathcal{P}_2(\mathbb{R}^d)$ and where the processes $\Big\{(X_t^{\xi,(i)})_{t\geq 0},\ 1 \leq i \leq N\Big\}$ are mutually independent copies of the process $(X_t^\xi)_{t\geq 0}$ defined on a copy $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Assume that b and σ are bounded and continuous functions. Assume that $(x,z) \mapsto \sigma(t,x,z)$ is η-Hölder continuous uniformly with respect to t, denoting by μ_N is the N-fold product measure of μ, that $a(t,x,\mu) := (\int \sigma(t,x,z)\,\mu_N(dz))(\int \sigma(t,x,z)\,\mu_N(dz))^*$ is uniformly elliptic, uniformly w.r.t to the variables t, x, μ.

Then, assumptions of Theorem 3.3 are fulfilled and the SDE (1.1) is well posed in the weak sense. If in addition $x \mapsto \sigma(t,x,z)$ is Lipschitz continuous uniformly with respect to t and z, assumptions of Corollary 3.4 are satisfied and strong well-posedness holds.

Again, an approximation argument that we do not detail here allows to handle the case of a bounded measurable drift function $(t,x,z) \mapsto b(t,x,z)$ for both weak and strong well-posedness.

Example 3.7. (Scalar interaction(s)) For some $N \geq 0$, for maps $\psi_1, \varphi_1 \cdots, \psi_N, \varphi_N : \mathbb{R}^d \to \mathbb{R}$, we consider the following non-linear SDE with coefficients $b : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^N \to \mathbb{R}^d \otimes \mathbb{R}^q$:

$$\begin{aligned} X_t^\xi = \xi &+ \int_0^t b\Big(s, X_s^\xi, \widetilde{\mathbb{E}}\big[\psi_1(\widetilde{X}_s^\xi)\big], \cdots, \widetilde{\mathbb{E}}\big[\psi_N(\widetilde{X}_s^\xi)\big]\Big)ds \\ &+ \int_0^t \sigma\Big(s, X_s^\xi, \widetilde{\mathbb{E}}\big[\varphi_1(\widetilde{X}_s^\xi)\big], \cdots, \widetilde{\mathbb{E}}\big[\varphi_N(\widetilde{X}_s^\xi)\big]\Big)dW_s \end{aligned} \tag{3.6}$$

with $[\xi] \in \mathcal{P}_2(\mathbb{R}^d)$ and where the process $(\widetilde{X}_t^\xi)_{t\geq 0}$ is a copy of $(X_t^\xi)_{t\geq 0}$ defined on a copy $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Assume that $\psi_1, \varphi_1, \cdots, \psi_N, \varphi_N$ are bounded continuous functions such that each φ_i, $i = 1, \cdots, N$, is η-Hölder continuous for some $\eta \in (0,1]$ and that b is a bounded and continuous function. Suppose that σ is a measurable map such that $a = \sigma\sigma^*$ is bounded and continuous and satisfies: $x \mapsto a(t,x,z)$ is η-Hölder continuous uniformly with respect to t and z; $z \mapsto a_{i,j}(t,x,z)$ is continuously differentiable with a bounded derivative for $1 \leq i,j \leq d$; $z \mapsto \partial_{z_k} a_{i,j}(t,x,z)$ is Lipschitz-continuous, uniformly with respect to the other variables for all $(i,j,k) \in \{1, \cdots, d\}^2 \times \{1, \cdots, N\}$; there exists $\lambda \geq 1$ such that for all $(t,u,x,z) \in \mathbb{R}_+ \times (\mathbb{R}^d)^2 \times \mathbb{R}^N$, $\lambda^{-1}|u|^2 \leq \langle a(t,x,z)u, u\rangle \leq \lambda|u|^2$.

Then, assumptions of Theorem 3.3 are fulfilled and the SDE (1.1) is well posed in the weak sense. If in addition $x \mapsto \sigma(t,x,z)$ is Lipschitz continuous, assumptions of Corollary 3.4 are satisfied and strong well posedness holds.

Finally, an approximation argument exposed in Remark 4.1 (for sake of simplicity in the one-dimensional setting $d = q = N = 1$) allows to establish weak uniqueness under the following weaker assumption: $(t, x, z) \mapsto b(t, x, z)$ and a are bounded and continuous functions, ψ_1 is bounded measurable, $x \mapsto \sigma(t, x, z)$ and φ_1 are η-Hölder, $z \mapsto \sigma(t, x, z)$ is continuously differentiable with a bounded derivative and σ^2 is uniformly elliptic.

Example 3.8. (Polynomials on the Wasserstein space) We consider the following scalar non-linear SDE

$$X_t^\xi = \xi + \int_0^t \prod_{i=1}^N \widetilde{\mathbb{E}}\Big[\psi_i\big(t, X_s^\xi, \widetilde{X}_s^\xi\big)\Big] ds + \int_0^t \prod_{i=1}^N \widetilde{\mathbb{E}}\Big[\varphi_i\big(t, X_s^\xi, \widetilde{X}_s^\xi\big)\Big] dW_s, \quad [\xi] \in \mathcal{P}_2(\mathbb{R}^d), \tag{3.7}$$

where $\psi_1, \varphi_1 \cdots, \psi_N, \varphi_N : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}$ and where the process $(\widetilde{X}_t^\xi)_{t\geq 0}$ is a copy of $(X_t^\xi)_{t\geq 0}$ defined on a copy $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Assume that $\psi_1, \varphi_1, \cdots, \psi_N, \varphi_N$ are bounded continuous functions and that the functions φ_i, $i = 1, \cdots, N$ are η-Hölder continuous in space (uniformly in time) for some $\eta \in (0, 1]$. Assume for sake of simplicity that there exists $\lambda > 0$ such that, for any $i \in \{1, \cdots, N\}$, for all $(t, x, z) \in \mathbb{R}_+ \times \mathbb{R}^2$, $\lambda < \varphi_i(t, x, z)$.

Then, assumptions of Theorem 3.3 are fulfilled and the SDE (1.1) is well posed in the weak sense. If in addition, each $x \mapsto \prod_{i=1}^N \int \varphi_i(t, x, z)\mu(dz)$ is Lipschitz continuous uniformly with respect to t and μ, assumptions of Corollary 3.4 are satisfied and strong well posedness holds.

A multi-dimensional version of (3.7) can be described as follows. We consider functions $\varphi_i : \mathbb{R}_+ \times (\mathbb{R}^d)^2 \to \mathbb{R}^{q_{i-1}} \otimes \mathbb{R}^{q_i}$, $i = 1, \cdots, N$, for some positive integers $q_0, \cdots, q_N$ satisfying $q_0 = d$ and $q_N = q$, where each $(x, z) \mapsto \varphi_i(t, x, z)$ is a bounded and η-Hölder continuous (uniformly with respect to t) function. For sake of simplicity, assume that each $a_i(t, x, \mu) := (\int \varphi_i(t, x, z)\mu(dz))(\int \varphi_i(t, x, z)\mu(dz))^*$, $i = 1, \cdots, N$, is uniformly elliptic so that $a(t, x, \mu) := (\prod_{i=1}^N \int \varphi_i(t, x, z)\mu(dz))(\prod_{i=1}^N \int \varphi_i(t, x, z)\mu(dz))^*$ is also uniformly elliptic.

Then, $(t, x, \mu) \mapsto b(t, x, \mu)$, with $b_j(t, x, \mu) := \prod_{i=1}^N \int \psi_{i,j}(t, x, z)\, \mu(dz)$, $j = 1, \cdots, d$, each $\psi_{i,j} : \mathbb{R}_+ \times (\mathbb{R}^d)^2 \to \mathbb{R}$ being a bounded and continuous function, and $(t, x, \mu) \mapsto \sigma(t, y, \mu) := \prod_{i=1}^N \int \varphi_i(t, x, z)\mu(dz)$ satisfy (**HC**), (**HR**), (**HE**) and $\mu \mapsto a(t, x, \mu)$ belongs to ($\mathbf{CS}_+$). Hence, the SDE (1.1) is well-posed in the weak sense. If in addition, each $x \mapsto \prod_{i=1}^N \int \varphi_i(t, x, z)\mu(dz)$ is Lipschitz continuous, then assumptions of Corollary (3.4) are satisfied and strong well posedness holds.

3.2. **On the density of the solution of the SDE** (1.1) **and its regularity properties.** Under the assumption of Theorem 3.3, by weak uniqueness, the law of the process $(X_t^{s,\xi})_{t\geq s}$ given by the unique solution to the SDE (1.1) starting from the initial distribution $\mu = [\xi]$ at time s only depends upon ξ through its law μ. Given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, it thus makes sense to consider $([X_t^{s,\xi}])_{t\geq s}$ as a function of μ without specifying the choice of the lifted random variable ξ that has μ as distribution. We then introduce, for any $x \in \mathbb{R}^d$, the following *decoupled stochastic flow* associated to the SDE (1.1)

$$X_t^{s,x,\mu} = x + \int_s^t b(r, X_r^{s,x,\mu}, [X_r^{s,\xi}])\, dr + \int_s^t \sigma(r, X_r^{s,x,\mu}, [X_r^{s,\xi}])\, dW_r. \tag{3.8}$$

We note that the previous equation is not a McKean-Vlasov SDE since the law appearing in the coefficients is not $[X_r^{s,x,\mu}]$ but rather $[X_r^{s,\xi}]$, that is, the law of the solution to the SDE (1.1) (starting at time s from the initial distribution μ) at time r. Under the assumptions of Theorem 3.3, the time-inhomogeneous martingale problem associated to the SDE (3.8) is well-posed, see e.g. Stroock and Varadhan [SV79]. In particular, weak existence and uniqueness in law holds for the SDE (3.8).

Moreover, from Friedman [Fri64], see also McKean and Singer [MS67], it follows that the transition density of the SDE (3.8) exists[2]. In particular, the random variable $X_t^{s,x,\mu}$ has a density that we denote by $z \mapsto p(\mu, s, t, x, z)$ which admits a representation in infinite series by means of the parametrix method that we now briefly describe. We refer the reader to [Fri64] or Konakov and Mammen [KM00] for more details. We now introduce the approximation process $(\widehat{X}_{t_2}^{t_1,x,\mu})_{t_2\geq t_1}$ obtained from the dynamics (3.8) by removing the drift and freezing the diffusion coefficient in space at a fixed point y, namely

$$\widehat{X}_{t_2}^{t_1,x,\mu} = x + \int_{t_1}^{t_2} \sigma(r, y, [X_r^{s,\xi}])\, dW_r. \tag{3.9}$$

[2]In [Fri64], it is proved that if $x \mapsto \bar{b}(r, x) = b(r, x, [X_r^{s,\xi}])$ is bounded and Hölder-continuous then the fundamental solution associated to the infinitesimal generator of (3.8) exists and is unique by means of the parametrix method. However, existence of the transition density as well as weak existence and weak uniqueness can be derived under the sole assumption that the drift is bounded and measurable and the diffusion matrix is uniformly elliptic and Hölder continuous.

The process $(\widehat{X}_{t_2}^{t_1,x,\mu})_{t_2\geq t_1}$ is a simple Gaussian process with transition density given explicitly by

$$\widehat{p}^{y}(\mu,s,t_1,t_2,x,z) := g\left(\int_{t_1}^{t_2} a(r,y,[X_r^{s,\xi}])\,dr, z-x\right).$$

To make the notation simpler, we will write $\widehat{p}(\mu,s,t_1,t_2,x,y) := \widehat{p}^{y}(\mu,s,t_1,t_2,x,y)$ and $\widehat{p}^{y}(\mu,s,t_2,x,z) = \widehat{p}^{y}(\mu,s,s,t_2,x,z)$. Note importantly that the variable y acts twice since it appears as a terminal point where the density is evaluated and also as the point where the diffusion coefficient is frozen. Note also that in what follows we need to separate between the starting time t_1 of the approximation process and the starting time s of the original McKean-Vlasov dynamics. We now introduce the two infinitesimal generators associated to the dynamics (3.8) and (3.9), namely

$$\mathcal{L}_{s,t}f(\mu,t,x) = \sum_{i=1}^{d} b_i(t,x,[X_t^{s,\xi}])\partial_{x_i}f(\mu,t,x) + \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,x,[X_t^{s,\xi}])\partial^2_{x_i,x_j}f(\mu,t,x),$$

$$\widehat{\mathcal{L}}_{s,t}f(\mu,t,x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{i,j}(t,y,[X_t^{s,\xi}])\partial^2_{x_i,x_j}f(\mu,t,x)$$

and define the parametrix kernel $\mathcal{H}$ for $(\mu,r,x,y)\in\mathcal{P}_2(\mathbb{R}^d)\times[s,t)\times(\mathbb{R}^d)^2$

$$\begin{aligned}\mathcal{H}(\mu,s,r,t,x,y) &:= (\mathcal{L}_{s,r}-\widehat{\mathcal{L}}_{s,r})\widehat{p}(\mu,s,r,t,x,y)\\ &= \sum_{i=1}^{d} b_i(r,x,[X_r^{s,\xi}])\partial_{x_i}\widehat{p}(\mu,s,r,t,x,y)\\ &\quad + \frac{1}{2}\sum_{i,j=1}^{d}(a_{i,j}(r,x,[X_r^{s,\xi}]) - a_{i,j}(r,y,[X_r^{s,\xi}]))\partial^2_{x_i,x_j}\widehat{p}(\mu,s,r,t,x,y).\end{aligned}$$

Now we define the following space-time convolution operator

$$(f\otimes g)(\mu,s,r,t,x,y) := \int_r^t\int_{\mathbb{R}^d} f(\mu,s,r,v,x,z)g(\mu,s,v,t,z,y)\,dz\,dv$$

and to simplify the notation we will write $(f\otimes g)(\mu,s,t,x,y) := (f\otimes g)(\mu,s,s,t,x,y)$, $\mathcal{H}(\mu,s,t,x,z) = \mathcal{H}(\mu,s,s,t,x,z)$ and proceed similarly for other maps. We also define $f\otimes\mathcal{H}^{(k)} = (f\otimes\mathcal{H}^{(k-1)})\otimes\mathcal{H}$ for $k\geq 1$ with the convention that $f\otimes\mathcal{H}^{(0)}\equiv f$. With these notations, the following parametrix expansion in infinite series of the transition $p(\mu,s,t,x,z)$ holds. Let $T>0$. For any $0\leq s<t\leq T$ and any $(\mu,x,y)\in\mathcal{P}_2(\mathbb{R}^d)\times(\mathbb{R}^d)^2$

$$p(\mu,s,t,x,y) = \sum_{k\geq 0}(\widehat{p}\otimes\mathcal{H}^{(k)})(\mu,s,t,x,y). \tag{3.10}$$

Moreover, the above infinite series converge absolutely and uniformly for $(\mu,x,y)\in\mathcal{P}_2(\mathbb{R}^d)\times(\mathbb{R}^d)^2$ and satisfies the following Gaussian upper-bound: for any $0\leq s<t\leq T$ and any $(\mu,x,y)\in\mathcal{P}_2(\mathbb{R}^d)\times(\mathbb{R}^d)^2$

$$p(\mu,s,t,x,y)\leq E_{\eta/2,1}(C(|b|_\infty+1))\,g(c(t-s),y-x) \tag{3.11}$$

where $C:=C(T,\lambda,\eta)$ and $c:=c(\lambda)$ are two positive constants. We refer to [MS67] for a proof based on Kolmogorov's backward and forward equations satisfied by p, see also Frikha [Fri17] for a proof based on probabilistic arguments.

Under the additional assumption that $x\mapsto b(t,x,\mu)$ is η-Hölder continuous (uniformly with respect to the variables (t,μ)), the series expansion (3.10) satisfied by the mapping $(s,x)\mapsto p(\mu,s,t,x,y)$ provides the unique solution of the Backward Kolmogorov equation, see e.g. [Fri64], namely:

$$\begin{cases}(\partial_s+\mathcal{L}_s)p(s,t,x,y)=0 & \text{for } (s,x,y)\in[0,t)\times(\mathbb{R}^d)^2\\ p(s,t,.,y)\to\delta_y(.) & \text{weakly as } s\uparrow t.\end{cases}$$

Moreover, the following Gaussian estimates hold:

$$|\partial_x^n p(\mu,s,t,x,y)|\leq\frac{C}{(t-s)^{\frac{n}{2}}}\,g(c(t-s),y-x),\quad n=0,1,2 \tag{3.12}$$

and

$$\forall\beta\in[0,\eta),\ |\partial_x^2 p(\mu,s,t,x_1,y)-\partial_x^2 p(\mu,s,t,x_2,y)|\leq C\frac{|x_1-x_2|^\beta}{(t-s)^{1+\frac{\beta}{2}}}\left[g(c(t-s),y-x_1)+g(c(t-s),y-x_2)\right] \tag{3.13}$$

for some positive constants $C := C(T, |b|_\infty, [b]_H, [a]_H, \lambda, \eta)$ and $c := c(\lambda)$. We refer again to [Fri64] for a proof of the above estimates.

A similar representation in infinite series is also valid for the density of the random variable $X_t^{s,\mu}$, denoted by $z \mapsto p(\mu, s, t, z)$, but we will not use it explicitly. Actually, we will make use of the following key relation

$$p(\mu, s, t, z) = \int_{\mathbb{R}^d} p(\mu, s, t, x, z)\, \mu(dx). \tag{3.14}$$

The representation in infinite series of $p(\mu, s, t, z)$ is thus obtained by integrating $x \mapsto p(\mu, s, t, x, z)$ against the initial distribution μ, in other words, $z \mapsto p(\mu, s, t, z)$ is the density of the image measure of the map $x \mapsto p(\mu, s, t, x, z)$ by the measure μ.

We introduce the following additional assumption on the coefficients:

(**HR**$_+$) The diffusion coefficient a satisfies assumption (**HR**) and the map $x \mapsto \widetilde{a}(t, x, y, \mu)$ is η-Hölder, with modulus still denoted by $[\widetilde{a}]_H$ for notational convenience, uniformly with respect to the other variables.

Moreover, the drift coefficient satisfies an assumption similar to (**HR**), namely:

– The function $\mathbb{R}^d \ni x \mapsto b(t, x, \mu) \in \mathbb{R}^d$ is uniformly η-Hölder continuous for some $\eta \in (0, 1]$, that is,

$$[b]_H := \sup_{t\geq 0,\, x\neq y,\, \mu\in\mathcal{P}_2(\mathbb{R}^d)} \frac{|b(t, x, \mu) - b(t, y, \mu)|}{|x - y|^\eta} < \infty.$$

– There exists a function $B : \mathbb{R}_+ \times (\mathbb{R}^d)^2 \times (\mathcal{P}_2(\mathbb{R}^d))^2 \to \mathbb{R}^N$ such that for every $(t, x, \nu, \mu) \in \mathbb{R}_+ \times \mathbb{R}^d \times (\mathcal{P}_2(\mathbb{R}^d))^2$ the function $z \mapsto B(t, x, z, \nu, \mu)$ is η-Hölder continuous, with modulus denoted by $[B]_H$, for some $\eta \in (0, 1]$, uniformly with respect to the other variables and satisfies the following estimate: for all $t \geq 0$, for all $(x, \nu, \mu) \in \mathbb{R}^d \times (\mathcal{P}_2(\mathbb{R}^d))^2$

$$|b(t, x, \nu) - b(t, x, \mu)| \leq \Big| \int B(t, x, z, \nu, \mu)(\nu - \mu)(dz) \Big|.$$

Our next result concerns the regularity properties of the two maps $(s, \mu) \mapsto p(\mu, s, t, z)$ and $(s, \mu, x) \mapsto p(\mu, s, t, x, z)$ and also important estimates on its derivatives. As mentioned above under the assumptions of Theorem 3.3 and (**HR**$_+$), $x \mapsto p(\mu, s, t, x, z)$ is two times continuously differentiable. In view of the relation (3.14), it thus suffices to investigate the smoothness of the map $(s, \mu, x) \mapsto p(\mu, s, t, x, z)$.

Theorem 3.9. Assume that (**HC**), (**HE**), (**HR**$_+$) and that both maps $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto b(t, x, \mu) \in \mathbb{R}^d$, $\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto a(t, x, \mu) \in \mathbb{R}^{d\times d}$ belong to (**CS**$_+$) (see also Remark 2.4).

Then, the mapping $[0, t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (s, x, \mu) \mapsto p(\mu, s, t, x, z)$ is in $\mathcal{C}^{1,2,2}([0, t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$. For any $T > 0$, there exist two positive constants $C := C(|b|_\infty, [b]_H, |\widetilde{b}|_\infty, [\widetilde{b}]_H, |a|_\infty, |\widetilde{a}|_\infty, [a]_H, [\widetilde{a}]_H, T)$, $c := c(\lambda)$, such that for any $(\mu, s, x, x', z, v, v') \in \mathcal{P}_2(\mathbb{R}^d) \times [0, t) \times (\mathbb{R}^d)^5$ and any $0 \leq s < t \leq T$

$$|\partial_v^n[\partial_\mu p(\mu, s, t, x, z)](v)| \leq \frac{C}{(t - s)^{\frac{1+n-\eta}{2}}} g(c(t - s), z - x),\ n = 0, 1, \tag{3.15}$$

$$|\partial_s p(\mu, s, t, x, z)| \leq \frac{C}{t - s} g(c(t - s), z - x), \tag{3.16}$$

$$\begin{aligned} &\forall \beta \in [0, \eta), \quad |\partial_s p(\mu, s, t, x, z)(v) - \partial_s p(\mu, s, t, x', z)(v)| \\ &\qquad \leq C \frac{|x - x'|^\beta}{(t - s)^{1+\frac{\beta}{2}}} \left\{ g(c(t - s), z - x) + g(c(t - s), z - x') \right\}, \end{aligned} \tag{3.17}$$

$$\begin{aligned} &|\partial_v^n[\partial_\mu p(\mu, s, t, x, z)](v) - \partial_v^n[\partial_\mu p(\mu, s, t, x', z)](v)| \\ &\qquad \leq C \frac{|x - x'|^\beta}{(t - s)^{\frac{1+n+\beta-\eta}{2}}} \left\{ g(c(t - s), z - x) + g(c(t - s), z - x') \right\}, \end{aligned} \tag{3.18}$$

where $\beta \in [0, 1]$ for $n = 0$ and $\beta \in [0, \eta)$ for $n = 1$,

$$\forall \beta \in [0, \eta), \quad |\partial_v[\partial_\mu p(\mu, s, t, x, z)](v) - \partial_v[\partial_\mu p(\mu, s, t, x, z)](v')| \leq C \frac{|v - v'|^\beta}{(t - s)^{1+\frac{\beta-\eta}{2}}} g(c(t - s), z - x), \tag{3.19}$$

There exist positive constants $C := C(|b|_\infty, [b]_H, |\widetilde{b}|_\infty, [\widetilde{b}]_H, |\widetilde{B}|_\infty, [\widetilde{B}]_H, |a|_\infty, [a]_H, |\widetilde{a}|_\infty, [\widetilde{a}]_H, |\widetilde{A}|_\infty, [\widetilde{A}]_H, T)$, $c := c(\lambda)$, such that for any $(\mu, \mu', s, x, z, v) \in (\mathcal{P}_2(\mathbb{R}^d))^2 \times [0,t) \times (\mathbb{R}^d)^3$,

$$|\partial_x^n p(\mu, s, t, x, z) - \partial_x^n p(\mu', s, t, x, z)](v)| \leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^{\frac{n+\beta-\eta}{2}}} g(c(t-s), z-x), \tag{3.20}$$

where $\beta \in [0,1]$ for $n = 0$, 1 and $\beta \in [0, \eta)$ for $n = 2$,

$$|\partial_v^n[\partial_\mu p(\mu, s, t, x, z)](v) - \partial_v^n[\partial_\mu p(\mu', s, t, x, z)](v)| \leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}} g(c(t-s), z-x) \tag{3.21}$$

where $\beta \in [0,1]$ for $n = 0$ and $\beta \in [0, \eta)$ for $n = 1$ and for all $(s_1, s_2) \in [0, t)$

$$\begin{aligned} &|\partial_x^n p(\mu, s_1, t, x, z) - \partial_x^n p(\mu, s_2, t, x, z)| \\ &\leq C \left\{ \frac{|s_1 - s_2|^\beta}{(t-s_1)^{\frac{n}{2}+\beta}} g(c(t-s_1), z-x) + \frac{|s_1 - s_2|^\beta}{(t-s_2)^{\frac{n}{2}+\beta}} g(c(t-s_2), z-x) \right\}, \end{aligned} \tag{3.22}$$

where $\beta \in [0,1]$ for $n = 0$, $\beta \in [0, \frac{1+\eta}{2})$ for $n = 1$ and $\beta \in [0, \frac{\eta}{2})$ for $n = 2$ and

$$\begin{aligned} &|\partial_v^n[\partial_\mu p(\mu, s_1, t, x, z)](v) - \partial_v[\partial_\mu p(\mu, s_2, t, x, z)](v)| \\ &\leq C \left\{ \frac{|s_1 - s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_1), z-x) + \frac{|s_1 - s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_2), z-x) \right\}, \end{aligned} \tag{3.23}$$

where $\beta \in [0, \frac{\eta}{2})$ for $n = 0$, 1.

3.3. **On the Cauchy problem related to the PDE** (1.2)**.** The previous regularity properties on the density of the random variables $X_t^{s,\xi}$ and $X_t^{s,x,\mu}$ allow us in turn to tackle the Cauchy problem in the strip $0 \leq t \leq T$ related to the PDE (1.2) on the Wasserstein space. The two real-valued maps f and h appearing in (1.2) will be assumed to satisfy the following conditions:

(**HST**)
- The two maps $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (t, z, \mu) \mapsto f(t, z, \mu)$ and $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (z, \mu) \mapsto h(z, \mu)$ are continuous, the two maps $\mu \mapsto f(t, x, \mu)$ and $\mu \mapsto h(x, \mu)$ being of class (**CS**). Moreover, the maps $[0,T] \times (\mathbb{R}^d)^2 \times \mathcal{P}_2(\mathbb{R}^d) \ni (t, x, z, \mu) \mapsto \widetilde{f}(t, x, z, \mu)$, $(\mathbb{R}^d)^2 \times \mathcal{P}_2(\mathbb{R}^d) \ni (x, z, \mu) \mapsto \widetilde{h}(x, z, \mu)$ are continuous.
- The two functions $z \mapsto f(t, z, \mu)$ and $z' \mapsto \widetilde{f}(t, z, z', \mu)$ are locally Hölder continuous with exponent η, uniformly with respect to the other variables.
- The maps f, h, $\widetilde{f}$ and $\widetilde{h}$ satisfy the following growth assumptions:

$$|f(t, z, \mu)| + |h(z, \mu)| \leq C \exp\left(\alpha \frac{|z|^2}{T}\right)(1 + M_2^q(\mu)), \tag{3.24}$$

$$|\widetilde{f}(t, z, z', \mu)| + |\widetilde{h}(z, z', \mu)| \leq C \exp\left(\alpha \frac{|z|^2}{T}\right)(1 + |z'|^2 + M_2^q(\mu)) \tag{3.25}$$

where $M_2(\mu) := \int_{\mathbb{R}^d} |x|^2 \mu(dx)$, for some positive constants $C := C(T)$, α and $q \geq 1$.

Theorem 3.10. Assume that the assumptions of Theorem 3.9 and that (**HST**) hold. Then, there exists a positive constant $c := c(\lambda)$ such that for any $\alpha < c$, the function U defined by

$$\begin{aligned} U(t, x, \mu) &:= \int_{\mathbb{R}^d} h(z, [X_T^{t,\xi}])\, p(\mu, t, T, x, z)\, dz - \int_t^T \int_{\mathbb{R}^d} f(s, z, [X_s^{t,\xi}])\, p(\mu, t, s, x, z)\, dz\, ds \\ &= \mathbb{E}\left[h(X_T^{t,x,\mu}, [X_T^{t,\xi}]) - \int_t^T f(s, X_s^{t,x,\mu}, [X_s^{t,\xi}])\, ds \right], \end{aligned} \tag{3.26}$$

where $\xi \in \mathbb{L}^2$ with $[\xi] = \mu$, is the unique solution of the Cauchy problem (1.2) (in the strip $0 \leq t \leq T$) and

$$|U(t, x, \mu)| \leq C \exp\left(k|x|^2\right)(1 + M_2^q(\mu)), \quad \text{for } (t, x, \mu) \in [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \tag{3.27}$$

where $C := C(T, |b|_\infty, |\widetilde{b}|_\infty, |\widetilde{a}|_\infty, \lambda, \eta)$, $k := k(T, \lambda, \alpha)$ are positive constants.

Moreover, U is unique among all of the classical solutions to the PDE (1.2) satisfying (2.4), T being replaced by any $T' \in [0, T)$, as well as the exponential growth assumption (3.27) and with h and f satisfying (3.24) and (3.25) for some positive constants k and α.

4. Well-posedness of the martingale problem

In this section, we investigate the well-posedness of the martingale problem of Definition 3.1 associated to the SDE (1.1). As mentioned in the introduction, the proof follows Stroock and Varadhan's perturbation argument, the underlying space being $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. We first investigate the simple case of non-linear SDEs obtained from (1.1) by removing the drift coefficient and by freezing the diffusion coefficient only in the space variable, importantly, the dependence in the measure variable remains unchanged. In a first part, we thus study the well-posedness of the following class of non-linear SDEs with dynamics

$$\bar{X}_t^{s,\xi} = \xi + \int_s^t \sigma(r, [\bar{X}_r^{s,\xi}])\, dW_r \tag{4.1}$$

where ξ is a random variable independent of W with law μ.

4.1. Well-posedness of the approximation process.

Unfortunately, there is no result in the literature that guarantees weak existence and uniqueness to the SDE (4.1) under the considered assumptions. We clarify this situation in the next Lemma.

Lemma 4.1. Under (**HR**) and (**HE**) there exists a unique weak solution to the SDE (4.1). Moreover, for any $(s,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, denoting by $\mathbb{P}_{s,x} \in \mathcal{P}(\mathcal{C}([0,\infty), \mathbb{R}^d))$ the unique solution to the associated martingale problem with initial distribution δ_x at time s, $(s,x) \mapsto \mathbb{P}_{s,x}(B)$ is measurable for any Borel subset B of the canonical space $\mathcal{C}([0,\infty), \mathbb{R}^d)$. Moreover, it is strong Markov.

Proof. We restrict our consideration to the case $s = 0$. The proof relies on the application of the Banach fixed point theorem to suitable map and complete metric space. For a fixed $T > 0$ and an initial condition $\mu \in \mathcal{P}(\mathbb{R}^d)$, we consider the following set

$$\begin{aligned}\mathscr{A}_{T,\mu} = \big\{ & P \in \mathcal{C}([0,T], \mathcal{P}(\mathbb{R}^d)) : P(0) = \mu, \\ & \forall t \in (0,T],\ P(t) \text{ is absolutely continuous w.r.t. the Lebesgue measure.}\big\}\end{aligned}$$

which is a complete metric space for the metric $d(P,P') := \sup_{t\in[0,T]} \int_{\mathbb{R}^d} |p - p'|(\mu, 0, t, z)dz$ where p and p' stands for measurable versions for the densities of P and P' respectively. We define the map $\mathscr{T} : \mathscr{A}_{T,\mu} \to \mathscr{A}_{T,\mu}$ which to a probability measure $Q \in \mathscr{A}_{T,\mu}$ associates the measure $\mathscr{T}(Q) \in \mathscr{A}_{T,\mu}$ induced by the process

$$\widehat{X}_t^{\xi} = \xi + \int_0^t \sigma(r, Q(r))\, dW_r, \quad t \in [0,T],$$

that is, $\mathscr{T}(Q)(t) = [\widehat{X}_t^{\xi}]$, $t \in [0,T]$. Note that any fixed point of $\mathscr{T}$ is a solution to the martingale problem. For P_1, $P_2 \in \mathscr{A}_{T,\mu}$, we consider the two following sequences of SDEs $(\bar{X}^{1,(m)})_{m\geq 0}$ and $(\bar{X}^{2,(m)})_{m\geq 0}$ with dynamics:

$$\forall m \geq 0, \forall t \in [0,T], \quad \bar{X}_t^{i,(m+1)} = \xi + \int_0^t \sigma(s, [\bar{X}_s^{i,(m)}])dW_s, \quad [\bar{X}_t^{i,(0)}] = P_i(t), \quad i = 1,2. \tag{4.2}$$

We denote by $P_i^{(m)} = (P_i^{(m)}(t))_{t\in[0,T]} \in \mathscr{A}_{T,\mu}$ the probability measure induced by $(\bar{X}_t^{i,(m)})_{t\in[0,T]}$. The density function of the random vector $\bar{X}_t^{i,(m+1)}$ is given by $z \mapsto p_i^{(m+1)}(\mu,0,t,z) = \int p_i^{(m+1)}(\mu,0,t,x,z)\, \mu(dx)$, with $p_i^{(m+1)}(\mu,0,t,x,z) = g(\int_0^t a(s,[\bar{X}_s^{i,(m)}])ds, z-x)$ and $a(s,[\bar{X}_s^{i,(m)}]) = (\sigma\sigma^*)(s,[\bar{X}_s^{i,(m)}])$. For $m \geq 1$, by the mean-value theorem, one has

$$\begin{aligned}&(p_1^{(m+1)} - p_2^{(m+1)})(\mu,0,t,x,z) \\ &\quad = \sum_{i,j=1}^d \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2}(H_2^{i,j}.g)\Big(\int_0^t (\lambda a(s,[\bar{X}_s^{1,(m)}]) + (1-\lambda)a(s,[\bar{X}_s^{2,(m)}])ds, z-x\Big) \\ &\qquad \times \int_0^t (a_{i,j}(s,[\bar{X}_s^{1,(m)}]) - a_{i,j}(s,[\bar{X}_s^{2,(m)}]))ds\, d\lambda.\end{aligned} \tag{4.3}$$

Using (**HR**)(ii) and then Fubini's theorem, for $m \geq 1$, we can bound the difference of the diffusion matrix between the two solutions as follows

$$
\begin{aligned}
|a_{i,j}(s,[\bar{X}_s^{1,(m)}]) - a_{i,j}(s,[\bar{X}_s^{2,(m)}])| &\leq |\int_{\mathbb{R}^d} A(s,z,[\bar{X}_s^{1,(m)}],[\bar{X}_s^{2,(m)}])(p_1^{(m)} - p_2^{(m)})(\mu,0,s,z)\,dz| \\
&\leq \Big|\int_{(\mathbb{R}^d)^2} A(s,z,[\bar{X}_s^{1,(m)}],[\bar{X}_s^{2,(m)}])(p_1^{(m)} - p_2^{(m)})(\mu,0,s,x,z)\,dz\,\mu(dx)\Big| \\
&= |\int_{(\mathbb{R}^d)^2} (A(s,z,[X_s^{1,(m)}],[X_s^{2,(m)}]) - A(s,x,[X_s^{1,(m)}],[X_s^{2,(m)}])) \\
&\quad \times (p_1^{(m)} - p_2^{(m)})(\mu,0,s,x,z)\,dz\,\mu(dx)| \\
&\leq \int_{(\mathbb{R}^d)^2} (|z-x|^\eta \wedge 1)\,|p_1^{(m)} - p_2^{(m)}|(\mu,0,s,x,z)\,dz\,\mu(dx) \qquad (4.4)
\end{aligned}
$$

and for $m = 0$, one gets $|a_{i,j}(s,[X_s^{1,(0)}]) - a_{i,j}(s,[X_s^{2,(0)}])| = |a_{i,j}(s,P_1(s)) - a_{i,j}(s,P_2(s))| \leq Cd(P_1,P_2)$, $1 \leq i,j \leq d$. Hence, combining (4.3), (4.4) together with (**HE**) and the space-time inequality (1.4), we obtain

$$
\begin{aligned}
&\int_{(\mathbb{R}^d)^2} (|z-x|^\eta \wedge 1)|p_1^{(m+1)} - p_2^{(m+1)}|(\mu,0,t,x,z)\,dz\,\mu(dx) \\
&\qquad \leq \frac{C}{t^{1-\frac{\eta}{2}}} \int_0^t ds \int_{(\mathbb{R}^d)^2} (|z-x|^\eta \wedge 1)\,|p_1^{(m)} - p_2^{(m)}|(\mu,0,s,x,z)\,dz\,\mu(dx)
\end{aligned}
$$

which in turn, by induction, easily yields

$$
\begin{aligned}
&\int_{(\mathbb{R}^d)^2} (|z-x|^\eta \wedge 1)|p_1^{(m)} - p_2^{(m)}|(\mu,0,t,x,z)\,dz\,\mu(dx) \\
&\qquad \leq \frac{C^m}{t^{1-\frac{\eta}{2}}} \int_{\Delta_m(t)} \prod_{k=1}^{m-1} s_k^{-1+\frac{\eta}{2}} \max_{i,j} |a_{i,j}(s_m,[X_{s_m}^{1,(0)}]) - a_{i,j}(s_m,[X_{s_m}^{2,(0)}])|\,ds_1 \cdots ds_m \\
&\qquad \leq \frac{(Ct^{\frac{\eta}{2}})^m}{\lceil 1 + (m-1)\frac{\eta}{2} \rceil !}\, d(P_1,P_2)
\end{aligned}
$$

where $\Delta_m(t) := \{(s_1,\cdots,s_m) \in [0,t]^p : s_{m+1} := 0 \leq s_p \leq s_{p-1} \leq \cdots \leq s_1 \leq t =: s_0\}$ for a fixed $t \geq 0$.

We now plug the previous estimate into (4.4) so that coming back to (4.3), we finally get

$$
d(P_1^{(m+1)},P_2^{(m+1)}) := \sup_{t\in[0,T]} \int |p_1^{(m+1)} - p_2^{(m+1)}|(\mu,0,t,z)\,dz \leq \frac{(CT^{\frac{\eta}{2}})^m}{\lceil 1 + m\frac{\eta}{2} \rceil !}\, d(P_1,P_2), \quad m \geq 1. \qquad (4.5)
$$

Since $\sum_{m\geq 0} \frac{(CT^{\frac{\eta}{2}})^m}{m!} < \infty$, the Banach fixed point theorem guarantees that the map $\mathscr{T}$ has a unique fixed point $\mathbb{P}^* \in \mathscr{A}_{T,\mu}$, for any $T > 0$. Hence, the martingale problem associated to (4.1) is well-posed on any compact interval $[0,T]$. Obviously, existence and uniqueness extends to $[0,\infty)$. The measurability and strong Markov properties follow as in [EK86].

□

4.2. **Regularity of the transition density of the approximation process.** In this section, we investigate the regularity properties of the transition density of the SDE (4.1), the regularity with respect to the measure variable being understood in the sense of Lions. Hence, from now on, we will always assume that the initial condition ξ has $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ as distribution. We first remark that, by weak uniqueness, the law of the process $(\bar{X}_t^{s,\xi})_{t\geq s}$ generated by the SDE (4.1) starting from ξ at time s only depends upon the law of ξ. Given $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, it thus makes sense to consider $([\bar{X}_t^{s,\xi}])_{t\geq s}$ as a function of μ without specifying the choice of the lifted random variable ξ that has μ as distribution. We then introduce, for any $x \in \mathbb{R}^d$, the following decoupled stochastic flow associated to the SDE (4.1)

$$
\bar{X}_t^{s,x,\mu} = x + \int_s^t \sigma(r,[\bar{X}_r^{s,\xi}])dW_r. \qquad (4.6)
$$

We denote by $z \mapsto \bar{p}(\mu,s,t,z)$ (resp. $z \mapsto \bar{p}(\mu,s,t,x,z)$) the density function of the random vector $\bar{X}_t^{s,\xi}$ (resp. $\bar{X}_t^{s,x,\mu}$) given by the unique solution of the SDE (4.1) (resp. the SDE (4.6)) taken at time t and starting from ξ with law $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ (resp. starting from $x \in \mathbb{R}^d$) at time s. Observe the following

relations

$$\bar{p}(\mu,s,t,z)=\int \bar{p}(\mu,s,t,x,z)\,\mu(dx),\quad \bar{p}(\mu,s,t,x,z)=g\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])\,dr,z-x\right).$$

Our next objective is to investigate the regularity properties of the two maps $(s,\mu)\mapsto\bar{p}(\mu,s,t,z)$ and $(s,x,\mu)\mapsto\bar{p}(\mu,s,t,x,z)$. Since $x\mapsto\bar{p}(\mu,s,t,x,z)\in\mathcal{C}^2(\mathbb{R}^d)$ and in view of the above relation between $\bar{p}(\mu,s,t,z)$ and $\bar{p}(\mu,s,t,x,z)$, it suffices to investigate the smoothness of the map $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\bar{p}(\mu,s,t,x,z)$.

Proposition 4.1. Assume that (**HR**), (**HE**) are satisfied for $(s,\mu)\mapsto a(s,\mu)$ and that $\mu\mapsto a(s,\mu)$ is in ($\mathbf{CS}_+$). Then, the map $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\bar{p}(\mu,s,t,x,z)$ is in $\mathcal{C}^{1,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.

Furthermore, for any $T>0$, there exist some positive constants $C:=C(|a|_\infty,|\widetilde{a}|_\infty,[\widetilde{a}]_H,T)$, $c:=c(\lambda)$ such that for all $(\mu,\mu',x,z,v,v')\in(\mathcal{P}_2(\mathbb{R}^d))^2\times(\mathbb{R}^d)^4$ and $0\le s<t\le T$, the following estimates hold

$$|\partial_v^n[\partial_\mu\bar{p}(\mu,s,t,x,z)](v)|\le\frac{C}{(t-s)^{\frac{1+n-\eta}{2}}}g(c(t-s),z-x),\ n=0,1, \tag{4.7}$$

$$|\partial_s\bar{p}(\mu,s,t,x,z)|\le\frac{C}{t-s}g(c(t-s),z-x), \tag{4.8}$$

$$\begin{aligned}\forall\beta\in[0,\eta),\quad &|\partial_v[\partial_\mu\bar{p}(\mu,s,t,x,z)](v)-\partial_v[\partial_\mu\bar{p}(\mu,s,t,x,z)](v')|\\ &\le C\frac{|v-v'|^\beta}{(t-s)^{1+\frac{\beta-\eta}{2}}}g(c(t-s),z-x).\end{aligned} \tag{4.9}$$

There exists some positive constant $C:=C(|a|_\infty,|\widetilde{a}|_\infty,[\widetilde{a}]_H,|\widetilde{A}|_\infty,[\widetilde{A}]_H,T)$ such that for any $(\mu,\mu',s,x,z,v)\in(\mathcal{P}_2(\mathbb{R}^d))^2\times[0,t)\times(\mathbb{R}^d)^3$ the following estimates hold

$$|\partial_v^n[\partial_\mu\bar{p}(\mu,s,t,x,z)](v)-\partial_v^n[\partial_\mu\bar{p}(\mu',s,t,x,z)](v)|\le C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}}g(c(t-s),z-x) \tag{4.10}$$

where $\beta\in[0,1]$ for $n=0$ and $\beta\in[0,\eta)$ for $n=1$ and for all $(s_1,s_2)\in[0,t)^2$,

$$\begin{aligned}&|\partial_v^n[\partial_\mu\bar{p}(\mu,s_1,t,x,z)](v)-\partial_v^n[\partial_\mu\bar{p}(\mu,s_2,t,x,z)](v)|\\ &\le C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}}\,g(c(t-s_1),z-x)+\frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}}\,g(c(t-s_2),z-x)\right\},\end{aligned} \tag{4.11}$$

where $\beta\in[0,\frac{1+\eta}{2})$ for $n=0$ and $\beta\in\left[0,\frac{\eta}{2}\right)$ for $n=1$,

$$\begin{aligned}\forall\beta\in[0,\frac{\eta}{2}),\quad &|\partial_s\bar{p}(\mu,s_1,t,x,z)]-\partial_s\bar{p}(\mu,s_2,t,x,z)|\\ &\le C\left\{\Big(\frac{|s_1-s_2|^\beta}{(t-s_1)^{1+\beta}}+\frac{|a(s_1,\mu)-a(s_2,\mu)|}{t-s_1}\Big)\,g(c(t-s_1),z-x)\right.\\ &\quad\left.+\Big(\frac{|s_1-s_2|^\beta}{(t-s_2)^{1+\beta}}+\frac{|a(s_1,\mu)-a(s_2,\mu)|}{t-s_2}\Big)\,g(c(t-s_2),z-x)\right\}.\end{aligned} \tag{4.12}$$

Proof. Let us introduce some notations. We identify any $d\times d$ matrix Γ as an $\mathbb{R}^{d^2}$-valued vector and for any $z\in\mathbb{R}^d$ define:

$$\begin{aligned}&f_z:\mathbb{R}^{d^2}\to\mathbb{R}\\ &\Gamma\mapsto f_z(\Gamma)=(2\pi)^{\frac{d}{2}}\det(\Gamma)^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\langle\Gamma^{-1}z,z\rangle\right).\end{aligned}$$

We denote by Df_z, its gradient seen as an $\mathbb{R}^{d^2}$-valued vector. Formally, using the fact that $\mu\mapsto a(t,\mu)$ belongs to (**CS**), the derivatives of the mapping $[0,t)\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,\mu)\mapsto\bar{p}(\mu,s,t,z)=$

$\int \bar{p}(\mu,s,t,x,z)\,\mu(dx)$, with $[\xi]=\mu$, satisfy the following relations

$$\begin{aligned}\partial_\mu \bar{p}(\mu,s,t,z)(v) &= \partial_\nu\Big[\int \bar{p}(\mu,s,t,x,z)\,\nu(dx)\Big]_{\nu=\mu}(v) + \partial_\nu\Big[\int \bar{p}(\nu,s,t,x,z)\,\mu(dx)\Big]_{\nu=\mu}(v)\\ &= (-H_1.g)\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr, z-v\right) + \int \partial_\mu \bar{p}(\mu,s,t,x,z)(v)\,\mu(dx),\\ \partial_\mu \bar{p}(\mu,s,t,x,z)(v) &= Df_{z-x}\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr\right).\int_s^t \partial_\mu a(r,[\bar{X}_r^{s,\xi}])(v)\,dr\\ &= Df_{z-x}\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr\right).\int_s^t \partial_\nu\Big[\int \widetilde{a}(r,y,[\bar{X}_r^{s,\xi}])\bar{p}(\nu,s,r,y)\,dy\Big]_{|\nu=\mu}(v)\,dr\\ \partial_s \bar{p}(\mu,s,t,x,z) &= Df_{z-x}\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr\right).\partial_s\Big[\int_s^t a(r,[\bar{X}_r^{s,\xi}])\,dr\Big]\\ &= Df_{z-x}\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr\right).\left(-a(s,\mu)+\int_s^t \partial_u\Big[\int \widetilde{a}(r,y,[\bar{X}_r^{s,\xi}])\bar{p}(\mu,u,r,y)\,dy\Big]_{|u=s}dr\right)\end{aligned}$$

where $\partial_\mu a(r,[\bar{X}_r^{s,\xi}])(v)=(\partial_\mu[a_{i,j}(r,[\bar{X}_r^{s,\xi}])](v))_k)_{1\le i,j,k\le d}$ and $\partial_\nu[\int \widetilde{a}(r,y,[\bar{X}_r^{s,\xi}])\bar{p}(\nu,s,r,y)\,dy]_{|\nu=\mu}(v)=((\partial_\nu[\int \widetilde{a}_{i,j}(r,y,[\bar{X}_r^{s,\xi}])\bar{p}(\nu,s,r,y)\,dy]_{|\nu=\mu}(v))_k)_{1\le i,j,k\le d}$ are seen as $d^2\times d$ matrices, for any fixed $v\in\mathbb{R}^d$. Similarly, by formally differentiating the previous equality with respect to v, the derivatives of the mappings $v\mapsto \partial_\mu\bar{p}(\mu,s,t,z)(v)$, $v\mapsto\partial_\mu\bar{p}(\mu,s,t,x,z)(v)$ satisfy the following identities

$$\begin{aligned}\partial_v[\partial_\mu\bar{p}(\mu,s,t,z)](v) &= (H_2.g)\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr, z-v\right)+\int \partial_v[\partial_\mu\bar{p}(\mu,s,t,x,z)](v)\,\mu(dx)\\ \partial_v[\partial_\mu\bar{p}(\mu,s,t,x,z)](v) &= Df_{z-x}\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr\right).\int_s^t \partial_v\partial_\nu[\int \widetilde{a}(r,y,[\bar{X}_r^{s,\xi}])\bar{p}(\nu,s,r,y)dy]_{|\nu=\mu}(v)\,dr\end{aligned}$$

where $\partial_v[\partial_\mu\bar{p}(\mu,s,t,z)](v)=(\partial_{v_j}(\partial_\mu\bar{p}(\mu,s,t,z)(v))_i)_{1\le i,j\le d}$. Of course, the previous computations are only formal since one does not know that the map $[0,r)\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,\mu)\mapsto a(r,[\bar{X}_r^{s,\xi}])\in\mathbb{R}^{d\times d}$ is in $\mathcal{C}^{1,2}([0,r)\times\mathcal{P}_2(\mathbb{R}^d))$, for $r\in[s,t]$. In order to make these formal differentiations rigorous, our strategy consists in using an approximation argument that we now expose.

Step 1: Construction of an approximation sequence and related estimates.

For a fixed $s\ge 0$ and an initial measure $P^0\in\mathcal{P}(\mathcal{C}([s,\infty),\mathbb{R}^d))$, we consider the sequence of probability measures $(\mathbb{P}^{(m)})_{m\ge0}$ on the canonical space $\mathcal{C}([s,\infty),\mathbb{R}^d)$ induced by the iterative scheme $\Big\{(\bar{X}_t^{s,\xi,(m)})_{t\ge s}, m\ge 0\Big\}$ with the following dynamics

$$m\ge0,\quad \bar{X}_t^{s,\xi,(m+1)}=\xi+\int_s^t\sigma(r,[\bar{X}_r^{s,\xi,(m)}])dW_r,\quad [\bar{X}_t^{s,\xi,(0)}]=P^0(t),\quad t\in[s,\infty). \tag{4.13}$$

Since the diffusion coefficient a is bounded and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, the sequence $(\mathbb{P}^{(m)})_{m\ge0}$ is tight. Relabelling the indices if necessary, we may assert that $(\mathbb{P}^{(m)})_{m\ge0}$ converges weakly to a probability measure $\mathbb{P}^\infty$. From standard arguments that we omit for sake of simplicity, that is, passing to the limit in the characterisation of the related martingale problem, we deduce that $\mathbb{P}^\infty$ is the (unique) probability measure $\mathbb{P}$ induced by the unique weak solution to the SDE (4.1). As a consequence, every convergent subsequence converges to the same limit $\mathbb{P}$. Hence, the original sequence $(\mathbb{P}^{(m)})_{m\ge1}$ converges weakly to $\mathbb{P}$.

We also introduce the sequence of decoupled stochastic flows $(\bar{X}_t^{s,x,\mu,(m)})_{t\ge s}$, $m\ge0$, associated to the recursive scheme (4.13), namely

$$m\ge0,\quad \bar{X}_t^{s,x,\mu,(m+1)}=x+\int_s^t\sigma(r,[\bar{X}_r^{s,\xi,(m)}])dW_r,\quad t\in[s,\infty). \tag{4.14}$$

Since $(\mathbb{P}^{(m)})_{m\ge0}$ converges weakly to $\mathbb{P}$ and $\mu\mapsto a(s,\mu)$ is continuous with respect to the weak topology, we deduce that for any fixed μ,s,t,x,z, the sequence of density functions $(p_m(\mu,s,t,x,z))_{m\ge1}$, $z\mapsto p_m(\mu,s,t,x,z)$ being the density of the random vector $\bar{X}_t^{s,x,\mu,(m)}$, converges to $\bar{p}(\mu,s,t,x,z)$ as $m\uparrow\infty$.

Then, again for a fixed (s,t,x,z) $(\mathbb{R}_+)^2\times(\mathbb{R}^d)^2$, with $0\le s<t\le T$, we consider the sequence of real-valued mappings $\Big\{\mathbb{L}^2\ni\xi\mapsto\tilde{p}_m(\xi,s,t,x,z), m\ge1\Big\}$ obtained by *lifting* the original sequence

$\big\{\mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto p_m(\mu,s,t,x,z), m \geq 1\big\}$ on the atomless probability space $(\Omega,\mathcal{A},\mathbb{P})$ [3].

We now prove by induction on m the following properties:

- Under (**HR**), (**HE**) and if $\mu \mapsto a(t,\mu)$ belongs to (**CS**), the continuous mapping $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \ni (s,x,\mu) \mapsto p_m(\mu,s,t,x,z)$ is in $\mathcal{C}^{0,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.

- Additionally, if $\mu \mapsto a(t,\mu)$ satisfies $(\mathbf{CS}_+)_1$, then there exist two positive constants $C := C(|a|_\infty,|\widetilde{a}|_\infty,[\widetilde{a}]_H,T)$, $c := c(\lambda)$, thus do not depending on m, such that for any $(\mu,s,x,x',z,v,v') \in \mathcal{P}_2(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^5$,

$$|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](v)| \leq \frac{C}{(t-s)^{\frac{1+n-\eta}{2}}} g(c(t-s),z-x),\ n=0,1, \tag{4.15}$$

$$\begin{aligned}&|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](v) - \partial_v^n[\partial_\mu p_m(\mu,s,t,x',z)](v)|\\ &\qquad\leq C\frac{|x-x'|^\beta}{(t-s)^{\frac{1+n+\beta-\eta}{2}}}\left\{g(c(t-s),z-x)+g(c(t-s),z-x')\right\},\end{aligned} \tag{4.16}$$

where $\beta\in[0,1]$ for $n=0$ and $\beta\in[0,\eta)$ for $n=1$,

$$\begin{aligned}\forall\beta\in[0,\eta),\quad &|\partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v) - \partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v')|\\ &\qquad\leq C\frac{|v-v'|^\beta}{(t-s)^{1+\frac{\beta-\eta}{2}}} g(c(t-s),z-x).\end{aligned} \tag{4.17}$$

- Additionally, if $\mu \mapsto a(t,\mu)$ satisfies $(\mathbf{CS}_+)_2$ so that $\mu \mapsto a(t,\mu)$ belongs to $(\mathbf{CS}_+)$, then there exist two positive constants $C := C(|a|_\infty,|\widetilde{a}|_\infty,[\widetilde{a}]_H,|\widetilde{A}|_\infty,[\widetilde{A}]_H,T)$, $c := c(\lambda)$, thus do not depending on m, such that for any $(\mu,\mu',s,x,z,v) \in (\mathcal{P}_2(\mathbb{R}^d))^2\times[0,t)\times(\mathbb{R}^d)^3$,

$$|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](v) - \partial_v^n[\partial_\mu p_m(\mu',s,t,x,z)](v)| \leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}} g(c(t-s),z-x), \tag{4.18}$$

where $\beta\in[0,1]$ for $n=0$ and $\beta\in[0,\eta)$ for $n=1$, and for all $(s_1,s_2)\in[0,t)^2$,

$$\begin{aligned}&|\partial_v^n[\partial_\mu p_m(\mu,s_1,t,x,z)](v) - \partial_v^n[\partial_\mu p_m(\mu,s_2,t,x,z)](v)|\\ &\leq C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_1),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_2),z-x)\right\},\end{aligned} \tag{4.19}$$

where $\beta\in[0,\frac{1+\eta}{2})$ for $n=0$ and $\beta\in[0,\frac{\eta}{2})$ for $n=1$.

For $m=1$, $\xi\mapsto\widetilde{p}_1(\xi,s,t,x,z) = g(\int_s^t a(r,P^0(r))dr, z-x)$ is clearly Fréchet differentiable with Fréchet derivative $D\widetilde{p}_1(\xi,s,t,x,z)=0$ so that $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d \ni (s,x,\mu,v)\mapsto\partial_\mu p_1(\mu,s,t,x,z)(v) := 0$, $\partial_v[\partial_\mu p_1(\mu,s,t,x,z)](v) = 0$ and the estimates (4.15) to (4.19) are clearly valid. Finally, since $\partial_x^n p_1(\mu,s,t,x,z) = -(H_n.g)(\int_s^t a(r,P^0(r))dr, z-x)$, $n\geq 1$, one also gets that the maps $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\partial_x p_1(\mu,s,t,x,z)$, $\partial_x^2 p_1(\mu,s,t,x,z)$ are continuous. We thus conclude that $(s,x,\mu)\mapsto p_1(\mu,s,t,x,z)\in\mathcal{C}^{0,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.

Let us assume that the induction hypothesis is valid at step m. We then remark that if $(s_n,\mu_n)_{n\geq1}$ is a sequence of $[0,t)\times\mathcal{P}_2(\mathbb{R}^d)$ satisfying $\lim_n|s_n-s| = \lim_n W_2(\mu_n,\mu) = 0$ for some $(s,\mu)\in[0,t)\times\mathcal{P}_2(\mathbb{R}^d)$,

[3]For sake of simplicity, the lifting procedure is done onto the same probability space that carries the unique weak solution $(\bar{X},W,\{\mathcal{F}_t\})$ to the SDE (4.1). Alternatively, one can enlarge the previous space and consider an arbitrary rich enough atomless probability space.

then for any Borel map h defined over $\mathbb{R}^d$ with at most quadratic growth one has

$$\begin{aligned}\int h(z)p_m(\mu_n,s_n,t,z)\,dz - \int h(z)p_m(\mu,s,t,x,z)\,dz \\ &= \int h(z)p_m(\mu_n,s_n,t,x,z)\,dz\,\mu_n(dx) - \int h(z)p_m(\mu,s,t,x,z)\,dz\,\mu(dx) \\ &= \int\int h(z)p_m(\mu_n,s_n,t,x,z)\,dz\,(\mu_n-\mu)(dx) \\ &\quad + \int\int h(z)[p_m(\mu_n,s_n,t,x,z)-p_m(\mu,s,t,x,z)]\,dz\,\mu(dx) := \mathrm{I}^n + \mathrm{II}^n.\end{aligned} \tag{4.20}$$

Let $R>0$. We introduce a smooth cutoff function $\eta_R:\mathbb{R}^d\to\mathbb{R}_+$ such that $\eta_R(x)=1$ for $|x|\leq R$ and $\eta_R(x)=0$ for $|x|\geq 2R$. We split the first term into two parts, namely $\mathrm{I}^n=\mathrm{I}^n_1+\mathrm{I}^n_2$ with $\mathrm{I}^n_1:=\int\int h(z)p_m(\mu_n,s_n,t,x,z)\,dz\,\eta_R(x)\,(\mu_n-\mu)(dx)$ and $\mathrm{I}^n_2:=\int\int h(z)p_m(\mu_n,s_n,t,x,z)\,dz\,(1-\eta_R(x))\,(\mu_n-\mu)(dx)$. Assuming without loss of generality that n is sufficiently large so that $|s_n-s|\leq (t-s)/2$, we remark that since $|\partial_x p_m(\mu_n,s_n,t,x,z)|\leq C(t-s_n)^{-\frac{1}{2}}g(c(t-s_n),z-x)\leq C(t-s)^{-\frac{1}{2}}g(c(t-s),z-x)$ for some positive constants C, c independent of n and m, the function $x\mapsto\int h(z)p_m(\mu_n,s_n,t,x,z)\,dz\,\eta_R(x)$ is Lipschitz continuous with modulus bounded uniformly w.r.t n. The first term I^n_1 thus goes to zero as n goes to infinity. Now, since h is of at most quadratic growth, $\limsup_n|\mathrm{I}^n_2|\leq C\limsup_n\left\{\int_{|x|\geq 2R}\mu_n(dx)+\int_{|x|\geq 2R}|x|^2\mu_n(dx)+\int_{|x|\geq 2R}\mu(dx)+\int_{|x|\geq 2R}|x|^2\mu(dx)\right\}\leq 2C\left\{\int_{|x|\geq 2R}\mu(dx)+\int_{|x|\geq 2R}|x|^2\mu(dx)\right\}$ so that choosing n and then R large enough, we see that I^n_2 can be made smaller as any prescribed $\varepsilon>0$. We thus deduce that I^n goes to zero as n goes to infinity. The last term II^n goes to zero as n goes to infinity by dominated convergence theorem and continuity of $(\mu,s)\mapsto p_m(\mu,s,r,x,z)$. We thus conclude that $W_2([\bar{X}^{s_n,\xi_n,(m)}_t],[\bar{X}^{s,\xi,(m)}_t])\to 0$ as $n\uparrow\infty$, where $[\xi_n]=\mu_n$ and $[\xi]=\mu$, so that $\lim_n a(t,[\bar{X}^{s_n,\xi_n,(m)}_t])=a(t,[\bar{X}^{s,\xi,(m)}_t])$ which in turn yields the continuity of the map $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto p_{m+1}(\mu,s,t,x,z)=g(\int_s^t a(r,[\bar{X}^{s,\xi,(m)}_r])dr,z-x)$. The continuity of the two maps $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\partial_x p_{m+1}(\mu,s,t,x,z)$, $\partial^2_x p_{m+1}(\mu,s,t,x,z)$ readily follows from a similar argument.

We next apply Lemma 2.1 to the map $\mu\mapsto a(r,\mu)$ along the flow of probability measures $([\bar{X}^{s,\xi,(m)}_r])_{s\geq r,\mu\in\mathcal{P}_2(\mathbb{R}^d)}$ to derive the differentiability of $\mu\mapsto a(r,[\bar{X}^{s,\xi,(m)}_r])$. Note that from the estimate (4.15) and the inequality $|\partial^n_x p_m(\mu,s,r,x,z)|\leq C(r-s)^{-\frac{n}{2}}g(c(r-s),z-x)$, $n=1,2$, the map $(x,\mu)\mapsto p_m(\mu,s,r,x,z)$ satisfies the condition (2.6) of Definition 2.3 (observe that we don't need the time regularity of the density). We thus deduce that $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto a(r,[\bar{X}^{s,\xi,(m)}_r])$ is partially $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ (see Chapter 5 of [CD18] for a definition of partial $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ regularity) and satisfies the key decomposition

$$\begin{aligned}\partial^n_v[\partial_\mu a(r,[\bar{X}^{s,\xi,(m)}_r])](v) &= \int(\widetilde{a}(r,y',[\bar{X}^{s,\xi,(m)}_r])-\widetilde{a}(r,v,[\bar{X}^{s,\xi,(m)}_r]))\partial^{1+n}_x p_m(\mu,s,r,v,y')\,dy' \\ &\quad + \int_{(\mathbb{R}^d)^2}(\widetilde{a}(r,y',[\bar{X}^{s,\xi,(m)}_r])-\widetilde{a}(r,x',[\bar{X}^{s,\xi,(m)}_r]))\partial^n_v[\partial_\mu p_m(\mu,s,r,x',y)](v)\,dy\,\mu(dx')\end{aligned} \tag{4.21}$$

Again note that the continuity of the map $[0,t)\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(s,\mu,v)\mapsto\partial^n_v[\partial_\mu a(r,[\bar{X}^{s,\xi,(m)}_r])](v)$ for $n=0,1$ follows from the continuity of the maps $[0,r)\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,\mu)\mapsto\widetilde{a}(r,z,[\bar{X}^{s,\xi,(m)}_r])$, $z\in\mathbb{R}^d$ and $(s,\mu,v)\mapsto\partial^{1+n}_x p_m(\mu,s,r,v,y')$, $\partial^n_v[\partial_\mu p_m(\mu,s,r,x',y')](v)$ as well as the estimate (4.15) and the inequality $|\partial^n_x p_m(\mu,s,r,v,y')|\leq C(r-s)^{-\frac{n}{2}}g(c(r-s),y'-v)$. Observe now that from the dynamics (4.14), one has $\xi\mapsto\tilde{p}_{m+1}(\xi,s,t,x,z)=g(\int_s^t a(r,[\bar{X}^{s,\xi,(m)}_r])dr,z-x)$. Here, there is a slight abuse of notations since we still denote by $a(r,[\bar{X}^{s,\xi,(m)}_r])$ the lift of the function $\mu\mapsto a(r,[\bar{X}^{s,\xi,(m)}_r])$. However, in the subsequent proof, we will not really work with it directly but rather on its Fréchet derivative. Hence, from the previous arguments, $\mathbb{L}^2\ni\xi\mapsto\widetilde{p}_{m+1}(\xi,s,t,x,z)$ is Fréchet differentiable with Fréchet derivative

$$\begin{aligned}D\widetilde{p}_{m+1}(\xi,s,t,x,z) &= Df_{z-x}\left(\int_s^t a(r,[\bar{X}^{s,\xi,(m)}_r])dr\right). \\ &\quad .\int_s^t\left\{\int_{\mathbb{R}^d}(\widetilde{a}(r,y',[\bar{X}^{s,\xi,(m)}_r])-\widetilde{a}(r,\xi,[\bar{X}^{s,\xi,(m)}_r]))\partial_x p_m(\mu,s,r,\xi,y')\,dy'\right. \\ &\quad \left.+\int_{(\mathbb{R}^d)^2}(\widetilde{a}(r,y',[\bar{X}^{s,\xi,(m)}_r])-\widetilde{a}(r,x',[\bar{X}^{s,\xi,(m)}_r]))\partial_\mu p_m(\mu,s,r,x',y')(\xi)\,dy'\,\mu(dx')\right\}dr\end{aligned} \tag{4.22}$$

so that the mapping $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(s,x,\mu,v)\mapsto\partial_\mu p_{m+1}(\mu,s,t,x,z)(v)$ is given by

$$\begin{aligned}\partial_\mu p_{m+1}(\mu,s,t,x,z)(v) &:= Df_{z-x}\left(\int_s^t a(r,[\bar X_r^{s,\xi,(m)}])dr\right)\\ &\cdot\int_s^t\left\{\int_{\mathbb{R}^d}(\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v,[\bar X_r^{s,\xi,(m)}]))\,\partial_x p_m(\mu,s,r,v,y')\,dy'\right.\\ &\left.+\int_{(\mathbb{R}^d)^2}(\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,x',[\bar X_r^{s,\xi,(m)}]))\,\partial_\mu p_m(\mu,s,r,x',y')(v)\,dy'\,\mu(dx')\right\}dr\end{aligned}\tag{4.23}$$

and it is globally continuous. Moreover, from (4.21), one deduces that the map $\mathbb{R}^d\ni v\mapsto\partial_\mu p_{m+1}(\mu,s,t,x,z)(v)$ is continuously differentiable with

$$\begin{aligned}&\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)\\ &:= Df_{z-x}\left(\int_s^t a(r,[\bar X_r^{s,\xi,(m)}])dr\right)\cdot\int_s^t\left\{\int_{\mathbb{R}^d}(\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v,[\bar X_r^{s,\xi,(m)}]))\partial_x^2 p_m(\mu,s,r,v,y')\,dy'\right.\\ &\left.+\int_{(\mathbb{R}^d)^2}(\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,x',[\bar X_r^{s,\xi,(m)}]))\partial_v[\partial_\mu p_m(\mu,s,r,x',y')](v)dy'\,\mu(dx')\right\}dr.\end{aligned}\tag{4.24}$$

Combining the previous expression with the induction hypothesis and the condition $(\mathbf{CS}_+)_1$ as well as the estimate (4.15) and the straightforward inequality $|\partial_x^2 p_m(\mu,s,r,v,y')|\le C(r-s)^{-1}g(c(r-s)y'-v)$, we also get that $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(s,x,\mu,v)\mapsto\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)$ is globally continuous. We thus conclude that $(s,x,\mu)\mapsto p_{m+1}(\mu,s,t,x,z)$ is in $\mathcal{C}^{0,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.

In order to establish the estimates (4.15), for $n=0,\,1$, at step $m+1$, we proceed as follows. Starting from the induction relations (4.23) and (4.24), since $\mu\mapsto a(s,\mu)$ is in (**CS**) and satisfies $(\mathbf{CS}_+)_1$, using (**HE**) we deduce that there exist some positive constants $C:=C(|a|_\infty,[\widetilde a]_H),c:=c(\lambda)$ independent of m such that

$$\begin{aligned}|\partial_v^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)| &\le\frac{C}{t-s}\left\{\int_s^t\Big[\int_{\mathbb{R}^d}\frac{(|y-v|^\eta\wedge1)}{(r-s)^{\frac{1+n}{2}}}\,g(c(r-s),y-v)\,dy\right.\\ &\left.+\int_{(\mathbb{R}^d)^2}(|y-x'|^\eta\wedge1)|\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y)](v)|\,dy\,\mu(dx')\Big]dr\right\}\\ &\times g(c(t-s),z-x)\end{aligned}\tag{4.25}$$

which in turn, by the space-time inequality (1.4), clearly implies the following relation

$$\begin{aligned}&\int_{(\mathbb{R}^d)^2}(|y-x|^\eta\wedge1)|\partial_v^n[\partial_\mu p_{m+1}(\mu,s,t,x,y)](v)|\,dy\,\mu(dx)\\ &\le\frac{C}{(t-s)^{1-\frac{\eta}{2}}}\left\{\int_s^t\frac{1}{(r-s)^{\frac{1+n-\eta}{2}}}dr+\int_s^t\int_{(\mathbb{R}^d)^2}(|y-x|^\eta\wedge1)|\partial_v^n[\partial_\mu p_m(\mu,s,r,x,y)](v)|\,dy\,\mu(dx)dr\right\}.\end{aligned}\tag{4.26}$$

Iterating the previous inequality, we deduce that there exists a constant C independent of m such that

$$\int_{(\mathbb{R}^d)^2}(|y'-x|^\eta\wedge1)|\partial_v^n[\partial_\mu p_{m+1}(\mu,s,t,x,y')](v)|\,dy'\,\mu(dx)\le\frac{C}{(t-s)^{\frac{1+n}{2}-\eta}}$$

which in turn, by (4.25), directly yields

$$|\partial_v^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)|\le\frac{C}{(t-s)^{\frac{1+n-\eta}{2}}}g(c(t-s),z-x),\quad n=0,\,1.$$

This concludes the proof of (4.15). Now, the estimate (4.16) easily follows from the previous estimate combined with the relations (4.23) and (4.24) so we omit its proof.

We now prove (4.17). From (4.24), we use the following decomposition

$$\begin{aligned}
&\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)-\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v')=Df_{z-x}\Big(\int_s^t a(r,y,[\bar X_r^{s,\xi,(m)}])dr\Big)\\
&.\int_s^t dr\,\Big\{\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v,[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v,y')\,dy'\\
&-\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v',[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v',y')\,dy'\\
&+\int_{(\mathbb{R}^d)^2}[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,x',[\bar X_r^{s,\xi,(m)}])](\partial_v[\partial_\mu p_m(\mu,s,r,x',y)](v)-\partial_v[\partial_\mu p_m(\mu,s,r,x',y)](v'))dy\mu(dx')\Big\}
\end{aligned}\tag{4.27}$$

which is valid for any $(v,v')\in(\mathbb{R}^d)^2$. First let us assume that $|v-v'|^2\le r-s$. From Fubini's theorem and the mean value theorem, we get

$$\begin{aligned}
&\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v,[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v,y')\,dy'\\
&\qquad-\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v',[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v',y')\,dy'\\
&=\int\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])[\partial_x^2 p_m(\mu,s,r,v,y')-\partial_x^2 p_m(\mu,s,r,v',y')]\,dy'\\
&=\int_0^1\int\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])\partial_x^3 p_m(\mu,s,r,\lambda v+(1-\lambda)v',y')(v-v')\,dy'\,d\lambda\\
&=\int_0^1\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,\lambda v+(1-\lambda)v',[\bar X_r^{s,\xi,(m)}])]\partial_x^3 p_m(\mu,s,r,\lambda v+(1-\lambda)v',y')(v-v')\,dy'\,d\lambda.
\end{aligned}$$

Using the direct bound $|\partial_x^3 p_m(\mu,s,r,x,y)|\le C(r-s)^{-\frac{3}{2}}g(c(r-s),y-x)$ and noting that for any point $\zeta\in(v,v')$, one has

$$\exp\Big(-\frac{|y-\zeta|^2}{c(r-s)}\Big)\le C\Big\{\exp\Big(-\frac{|y-v|^2}{c(r-s)}\Big)+\exp\Big(-\frac{|y-v'|^2}{c(r-s)}\Big)\Big\}$$

from the space-time inequality (1.4) and condition $(\mathbf{CS}_+)_1$, we deduce

$$\begin{aligned}
&\Big|\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v,[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v,y')\,dy'\\
&\qquad-\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v',[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v',y')\,dy'\Big|\\
&\le C\frac{|v-v'|}{(r-s)^{\frac{3-\eta}{2}}}\\
&\le C\frac{|v-v'|^\beta}{(r-s)^{1+\frac{\beta-\eta}{2}}}
\end{aligned}$$

for any $\beta\in[0,1]$. If $|v-v'|^2>r-s$, then from (1.5), condition $(\mathbf{CS}_+)_1$ and (1.4), we directly get

$$\begin{aligned}
\Big|\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v,[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v,y')\,dy'\Big|&\le C\int\frac{|v-v'|^\beta}{(r-s)^{1+\frac{\beta-\eta}{2}}}g(c(r-s),y'-v)\,dy'\\
&\le\frac{|v-v'|^\beta}{(r-s)^{1+\frac{\beta-\eta}{2}}}
\end{aligned}$$

and similarly

$$\Big|\int[\widetilde a(r,y',[\bar X_r^{s,\xi,(m)}])-\widetilde a(r,v',[\bar X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v',y')\,dy'\Big|\le C\frac{|v-v'|^\beta}{(r-s)^{1+\frac{\beta-\eta}{2}}}.$$

Hence, for all $(v,v')\in(\mathbb{R}^d)^2$ and for all $\beta\in[0,1]$, one has

$$\begin{aligned}&\Big|\int[\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,v,[\bar{X}_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v,y')\,dy'\\&\qquad-\int[\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,v',[\bar{X}_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,v',y')\,dy'\Big|\\&\le C\frac{|v-v'|^\beta}{(r-s)^{1+\frac{\beta-\eta}{2}}}.\end{aligned}$$

We now plug the previous estimates into (4.27) so that we obtain

$$\begin{aligned}&|\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)-\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v')|\le\frac{C}{t-s}\Big\{\int_s^t\Big[\frac{|v-v'|^\beta}{(r-s)^{1+\frac{(\beta-\eta)}{2}}}\\&+\int_{(\mathbb{R}^d)^2}(|y'-x'|^\eta\wedge1)|\partial_v[\partial_\mu p_m(\mu,s,r,x',y)](v)-\partial_v[\partial_\mu p_m(\mu,s,r,x',y')](v')|\,dy'\,\mu(dx')\Big]dr\Big\}\,g(c(t-s),z-x)\end{aligned}$$

which in turn implies

$$\begin{aligned}&\int_{(\mathbb{R}^d)^2}(|y-x|^\eta\wedge1)|\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,y)](v)-\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,y)](v')|\,dy\,\mu(dx)\\&\quad\le\frac{C}{(t-s)^{1-\frac{\eta}{2}}}\Big\{\int_s^t\Big[\frac{|v-v'|^\beta}{(r-s)^{1+\frac{(\beta-\eta)}{2}}}\\&\qquad+\int_{(\mathbb{R}^d)^2}(|y-x|^\eta\wedge1)|\partial_v[\partial_\mu p_m(\mu,s,r,x,y)](v)-\partial_v[\partial_\mu p_m(\mu,s,r,x,y)](v'))\,dy\,\mu(dx)|\Big]dr\Big\}.\end{aligned}$$

From the previous relation, using similar arguments as those employed for the previous estimates, we obtain (4.17).

It now remains to prove (4.18) and (4.19) under the additional assumption $(\mathbf{CS}_+)_2$, or in other words, under the assumption that the map $\mu\mapsto a(s,\mu)$ is in $(\mathbf{CS}_+)$. Since the arguments are quite similar for both estimates, we prove (4.18) and shall be brief for (4.19). From now on, we assume that $\mu\mapsto a(s,\mu)$ belongs to $(\mathbf{CS}_+)$. In order to establish (4.18) we make use of the following decomposition:

$$\partial_v^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)-\partial_v^n[\partial_\mu p_{m+1}(\mu',s,t,x,z)](v)=\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}$$

with

$$\begin{aligned}\mathrm{I}&:=\Big\{Df_{z-x}\Big(\int_s^t a(r,[\bar{X}_r^{s,\xi,(m)}])dr\Big)-Df_{z-x}\Big(\int_s^t a(r,[\bar{X}_r^{s,\xi',(m)}])dr\Big)\Big\}\\&\quad.\int_s^t\Big\{\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,v,[\bar{X}_r^{s,\xi,(m)}]))\partial_x^{1+n}p_m(\mu,s,r,v,y')\,dy'\\&\qquad+\int\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,x',[\bar{X}_r^{s,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)\,dy'\,\mu(dx')\Big\}\,dr,\\\mathrm{II}&:=Df_{z-x}\Big(\int_s^t a(r,[\bar{X}_r^{s,\xi',(m)}])dr\Big).\int_s^t\Big\{\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,y',[\bar{X}_r^{s,\xi',(m)}]))\partial_x^{1+n}p_m(\mu,s,r,v,y')\,dy'\\&\qquad+\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi',(m)}])-\widetilde{a}(r,v,[\bar{X}_r^{s,\xi',(m)}]))(\partial_x^{1+n}p_m(\mu,s,r,v,y')-\partial_x^{1+n}p_m(\mu',s,r,v,y'))\,dy'\Big\}\,dr,\\\mathrm{III}&:=Df_{z-x}\Big(\int_s^t a(r,[\bar{X}_r^{s,\xi',(m)}])dr\Big).\int_s^t\Big\{\int\int\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)\,dy'\,(\mu-\mu')(dx')\\&\qquad+\int\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,y',[\bar{X}_r^{s,\xi',(m)}]))\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)\,dy'\,\mu'(dx')\Big\}\,dr,\\\mathrm{IV}&:=Df_{z-x}\Big(\int_s^t a(r,[\bar{X}_r^{s,\xi',(m)}])dr\Big).\int_s^t\int\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi,(m)}])-\widetilde{a}(r,y',[\bar{X}_r^{s,\xi',(m)}]))\\&\quad.(\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)-\partial_v^n[\partial_\mu p_m(\mu',s,r,x',y')](v))\,dy'\,\mu'(dx')\,dr\end{aligned}$$

and will prove the following estimates

$$
\begin{aligned}
|\mathrm{I}| &\leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta}{2}-\eta}}g(c(t-s),z-x), \quad \forall \beta\in[0,1],\\
|\mathrm{II}| &\leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}}g(c(t-s),z-x), \text{ where } \beta\in[0,\frac{1+\eta}{2}), \text{ if } n=0, \text{and } \beta\in[0,\eta), \text{ if } n=1,\\
|\mathrm{III}| &\leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta}{2}-\eta}}g(c(t-s),z-x), \text{ where } \beta\in[0,1], \text{ if } n=0, \text{and } \beta\in[0,2\eta), \text{ if } n=1,\\
|\mathrm{IV}| &\leq C\frac{1}{t-s}\Big(\int_s^t\int\int(|y'-x'|^\eta\wedge 1)\\
&\quad\times|\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)-\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)|\,dy'\,\mu'(dx')\,dr\Big)\\
&\quad\times g(c(t-s),z-x).
\end{aligned}
$$

Before that, note that the previous estimates yield

$$
\begin{aligned}
&|\partial_v^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)-\partial_v^n[\partial_\mu p_{m+1}(\mu',s,t,x,z)](v)|\leq C\Big[\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}}+\frac{1}{t-s}\\
&\times\Big(\int\int(|y'-x'|^\eta\wedge1)|\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)-\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)|\,dy'\,\mu'(dx')\Big)\Big]g(c(t-s),z-x)
\end{aligned}
$$

which in turn, by an induction argument similar to the previous ones, which is omitted, implies (4.18). We now prove the announced estimates on I, II, III, IV.

∘ *Estimate on* I*:*

From (**HR**)(ii), for any $\beta\in[0,1]$, we get the following intermediate estimate

$$
\begin{aligned}
&\Big|a(r,[\bar X_r^{s,\xi,(m)}])-a(r,[\bar X_r^{s,\xi',(m)}])\Big|\\
&\qquad\leq\Big|\int A(r,y',[\bar X_r^{s,\xi,(m)}],[\bar X_r^{s,\xi',(m)}])(p_m(\mu,s,r,y')-p_m(\mu',s,r,y'))\,dy'\Big| \qquad (4.28)\\
&\qquad\leq\Big|\int\int A(r,y',[\bar X_r^{s,\xi,(m)}],[\bar X_r^{s,\xi',(m)}])p_m(\mu,s,r,x',y')\,dy'\,(\mu-\mu')(dx')\Big|\\
&\qquad\quad+\Big|\int\int\Big\{A(r,y',[\bar X_r^{s,\xi,(m)}],[\bar X_r^{s,\xi',(m)}])-A(r,x',[\bar X_r^{s,\xi,(m)}],[\bar X_r^{s,\xi',(m)}])\Big\}\\
&\qquad\qquad\times(p_m(\mu,s,r,x',y')-p_m(\mu',s,r,x',y'))\,dy'\,\mu'(dx')\Big|\\
&\qquad\leq\Big|\int\int A(r,y',[\bar X_r^{s,\xi,(m)}],[\bar X_r^{s,\xi',(m)}])p_m(\mu,s,r,x',y')\,dy'\,(\mu-\mu')(dx')\Big|\\
&\qquad\quad+\int\int(|y'-x'|^\eta\wedge1)|p_m(\mu,s,r,x',y')-p_m(\mu',s,r,x',y')|\,dy'\,\mu'(dx')\\
&\qquad\leq\frac{W_2^\beta(\mu,\mu')}{(r-s)^{\frac{\beta-\eta}{2}}}+C\int\int(|y'-x'|^\eta\wedge1)|p_m(\mu,s,r,x',y')-p_m(\mu',s,r,x',y')|\,dy'\,\mu'(dx') \qquad (4.29)
\end{aligned}
$$

where we used the fact that $x\mapsto\int A(r,y,[\bar X_r^{s,\xi,(m)}],[\bar X_r^{s,\xi',(m)}])p_m(\mu,s,r,x,y)\,dy$ is β-Hölder with modulus bounded by $C(r-s)^{\frac{\eta-\beta}{2}}$, for any $\beta\in[0,1]$, for some positive constant $C:=C(a)$. Combining the previous computation with the mean value theorem and (1.5), we obtain

$$
\begin{aligned}
&|p_{m+1}(\mu,s,t,x,z)-p_{m+1}(\mu',s,t,x,z)|\\
&\qquad\leq\frac{C}{t-s}g(c(t-s),z-x)\int_s^t|a(r,[\bar X_r^{s,\xi,(m)}])-a(r,[\bar X_r^{s,\xi',(m)}])|dr\\
&\qquad\leq\frac{C}{t-s}g(c(t-s),z-x)\Big((t-s)^{1+\frac{\eta-\beta}{2}}W_2^\beta(\mu,\mu')\\
&\qquad+\int_s^t\int_{(\mathbb{R}^d)^2}(|y'-x'|^\eta\wedge1)|p_m(\mu,s,r,x',y')-p_m(\mu',s,r,x',y')|\,dy'\,\mu'(dx')\Big)
\end{aligned}
$$

and, employing similar arguments as those previously used, we deduce that there exists a positive constant $C := C(T, \lambda)$ such that

$$\forall \beta \in [0,1], \qquad |p_{m+1}(\mu, s, t, x, z) - p_{m+1}(\mu', s, t, x, z)| \leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^{\frac{\beta-\eta}{2}}} g(c(t-s), z-x). \tag{4.30}$$

More generally, in a completely analogous manner, we also obtain the estimates: for any $m \geq 0$

$$\forall \beta \in [0,1], |\partial_x^n p_{m+1}(\mu, s, t, x, z) - \partial_x^n p_{m+1}(\mu', s, t, x, z)| \leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^{\frac{n+\beta-\eta}{2}}} g(c(t-s), z-x), \quad n = 0, 1, 2. \tag{4.31}$$

We now come back to I. Plugging (4.30) into (4.29), we finally deduce

$$\left| a(r, [\bar{X}_r^{s,\xi,(m)}]) - a(r, [\bar{X}_r^{s,\xi',(m)}]) \right| \leq C \frac{W_2^\beta(\mu, \mu')}{(r-s)^{\frac{\beta-\eta}{2}}}$$

so that, using the mean-value theorem as well as $(\mathbf{CS}_+)_1$, (1.5), (4.15) and the space-time inequality (1.4), we get

$$\begin{aligned} \forall \beta \in [0,1], \quad |\mathrm{I}| &\leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^2} \left(\int_s^t \frac{1}{(r-s)^{\frac{\beta-\eta}{2}}} dr \right) \left(\int_s^t \frac{1}{(r-s)^{\frac{1+n-\eta}{2}}} dr \right) g(c(t-s), z-x) \\ &\leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^{\frac{1+n+\beta}{2}-\eta}} g(c(t-s), z-x). \end{aligned}$$

∘ *Estimate on* II*:*

Following similar lines of reasonings as those employed to establish (4.29), namely using the fact that $\mu \mapsto a(r, \mu)$ satisfies $(\mathbf{CS}_+)_2$ instead of $(\mathbf{HR})$(ii), one gets

$$\begin{aligned} \left| \widetilde{a}(r, y, [\bar{X}_r^{s,\xi,(m)}]) - \widetilde{a}(r, y, [\bar{X}_r^{s,\xi',(m)}]) \right| &\leq \left| \int \widetilde{A}(r, y, z, [\bar{X}_r^{s,\xi,(m)}], [\bar{X}_r^{s,\xi',(m)}]) (p_m(\mu, s, r, z) - p_m(\mu', s, r, z))\, dz \right| \\ &\leq C \frac{W_2^\beta(\mu, \mu')}{(r-s)^{\frac{\beta-\eta}{2}}} \end{aligned} \tag{4.32}$$

which in turn, by (1.5), directly yields

$$\left| \int (\widetilde{a}(r, y, [\bar{X}_r^{s,\xi,(m)}]) - \widetilde{a}(r, y, [\bar{X}_r^{s,\xi',(m)}])) \partial_x^{1+n} p_m(\mu, s, r, v, z)\, dz \right| \leq C \frac{W_2^\beta(\mu, \mu')}{(r-s)^{\frac{1+n+\beta-\eta}{2}}}.$$

Moreover, from (4.31), the η-Hölder regularity of $y \mapsto \widetilde{a}(r, y, \mu)$ and the space-time inequality (1.4)

$$\left| \int (\widetilde{a}(r, y, [\bar{X}_r^{s,\xi',(m)}]) - \widetilde{a}(r, v, [\bar{X}_r^{s,\xi',(m)}])) (\partial_x^{1+n} p_m(\mu, s, r, v, y') - \partial_x^{1+n} p_m(\mu', s, r, v, y'))\, dy' \right| \leq C \frac{W_2^\beta(\mu, \mu')}{(r-s)^{\frac{1+n+\beta}{2}-\eta}}$$

so that combining the two previous bounds finally yield

$$|\mathrm{II}| \leq C \frac{W_2^\beta(\mu, \mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}} g(c(t-s), z-x).$$

where $\beta \in [0, \frac{1+\eta}{2})$ if $n = 0$ and $\beta \in [0, \eta)$ if $n = 1$.

∘ *Estimate on* III*:*

From the relation (4.24), the mean-value theorem and using the fact that $y \mapsto \widetilde{a}(r, y, \mu)$ is η-Hölder, it follows that $x \mapsto \int \widetilde{a}(r, y, [\bar{X}_r^{s,\xi,(m)}]) \partial_v^n [\partial_\mu p_m(\mu, s, r, x, y)](v)\, dy$ is β-Hölder continuous with a modulus bounded by $C(r-s)^{-\frac{(1+n+\beta)}{2}+\eta}$, for any $\beta \in [0,1]$, for some positive constant $C := C(|\widetilde{a}|_\infty, [\widetilde{a}]_H)$. Hence, one has

$$\left| \int \int \widetilde{a}(r, y, [\bar{X}_r^{s,\xi,(m)}]) \partial_v^n [\partial_\mu p_m(\mu, s, r, x, y)](v)\, dy\, (\mu - \mu')(dx) \right| \leq C \frac{W_2^\beta(\mu, \mu')}{(r-s)^{\frac{1+n+\beta}{2}-\eta}}.$$

From (4.32) and (4.15), we get

$$\left| \int \int (\widetilde{a}(r, y, [\bar{X}_r^{s,\xi,(m)}]) - \widetilde{a}(r, y, [\bar{X}_r^{s,\xi',(m)}])) \partial_v^n [\partial_\mu p_m(\mu, s, r, x, y)](v)\, dy\, \mu'(dx) \right| \leq C \frac{W_2^\beta(\mu, \mu')}{(r-s)^{\frac{1+n+\beta}{2}-\eta}}.$$

Consequently, combining the two previous bounds, we obtain

$$|\mathrm{III}| \leq C\frac{W_2^\beta(\mu,\mu')}{t-s}\left(\int_s^t \frac{1}{(r-s)^{\frac{1+n+\beta}{2}-\eta}}dr\right) g(c(t-s),z-x) \leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta}{2}-\eta}} g(c(t-s),z-x)$$

where $\beta \in [0,1]$ if $n=0$ and $\beta \in [0,2\eta)$ if $n=1$.

∘ *Estimate on* IV*:*

For the last term, using that $y \mapsto \widetilde{a}(r,y,\mu)$ is η-Hölder, we get

$$|\mathrm{IV}| \leq \frac{C}{t-s}\left(\int\int (|y'-x'|^\eta \wedge 1)|\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v) - \partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)|\, dy'\, \mu'(dx')\right) \times g(c(t-s),z-x).$$

This last inequality completes the proof of the announced estimates on I, II, III, IV.

The estimate (4.19) is proved by following similar lines of reasonings. We only sketch its proof and omit some technical details. Using (**HR**), the mean value theorem and the space-time inequality (1.4), we get

$$|p_{m+1}(\mu,s_1,t,x,z) - p_{m+1}(\mu,s_2,t,x,z)| \leq \int_0^1 \frac{C}{(t-(\lambda s_1+(1-\lambda)s_2))} g(c(t-(\lambda s_1+(1-\lambda)s_2)),z-x)d\lambda$$

$$\times \left(|s_1-s_2| + \int_{s_1\vee s_2}^t \int\int (|y'-x'|^\eta \wedge 1)|p_m(\mu,s_1,r,x',y') - p_m(\mu,s_2,r,x',y')|\, dy'\, \mu(dx')\, dr\right) \tag{4.33}$$

so that

$$\begin{aligned}
&\int\int (|z-x|^\eta \wedge 1)|p_{m+1}(\mu,s_1,t,x,z) - p_{m+1}(\mu,s_2,t,x,z)|\, dz\, \mu(dx) \\
&\quad \leq \Big(C\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\beta-\frac{\eta}{2}}} \\
&\quad + \frac{C}{(t-s_1\vee s_2)^{1-\frac{\eta}{2}}}\int_{s_1\vee s_2}^t \int\int (|z-x|^\eta \wedge 1)|p_m(\mu,s_1,r,x,z) - p_m(\mu,s_2,r,x,z)|dz\mu(dx)dr\Big)
\end{aligned}$$

if $|s_1-s_2| \leq t-s_1\vee s_2$. Hence, by an induction argument that we omit, we obtain

$$\forall \beta \in [0,1], \quad \int\int (|z-x|^\eta \wedge 1)\, |p_{m+1}(\mu,s_1,t,x,z) - p_{m+1}(\mu,s_2,t,x,z)|\, dz\, \mu(dx) \leq C\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\beta-\frac{\eta}{2}}} \tag{4.34}$$

for some positive constant $C := C(T,\lambda,[A]_H)$ if $|s_1-s_2| \leq t-s_1\vee s_2$. Now, if $|s_1-s_2| \geq t-s_1\vee s_2$, (4.34) easily follows from the space-time inequality (1.4). More generally, from a completely analogous argument, we obtain

$$\forall x \in \mathbb{R}^d, \quad \int (|z-x|^\eta \wedge 1)\, |\partial_x^n p_m(\mu,s_1,t,x,z) - \partial_x^n p_m(\mu,s_2,t,x,z)|\, dz \leq C\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{n-\eta}{2}+\beta}} \tag{4.35}$$

which directly implies

$$\int\int (|z-x|^\eta \wedge 1)\, |\partial_x^n p_m(\mu,s_1,t,x,z) - \partial_x^n p_m(\mu,s_2,t,x,z)|\, dz\, \mu(dx) \leq C\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{n-\eta}{2}+\beta}} \tag{4.36}$$

where $\beta \in [0,1]$ for $n=0$, $\beta \in [0,\frac{1+\eta}{2})$ for $n=1$, and $\beta \in [0,\frac{\eta}{2})$ for $n=2$. Hence, plugging (4.34) into (4.33), we get

$$\begin{aligned}
&\forall \beta \in [0,1], \quad |p_m(\mu,s_1,t,x,z) - p_m(\mu,s_2,t,x,z)| \\
&\qquad \leq C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^\beta} g(c(t-s_1),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^\beta} g(c(t-s_2),z-x)\right\}.
\end{aligned}$$

if $|s_1-s_2| \leq t-s_1\vee s_2$. Otherwise, if $|s_1-s_2| \geq t-s_1\vee s_2$, we directly get

$$\begin{aligned}
&\forall \beta \in [0,1], |p_m(\mu,s_1,t,x,z) - p_m(\mu,s_2,t,x,z)| \\
&\qquad \leq \frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^\beta} g(c(t-s_1\vee s_2),z-x) + \frac{(t-s_1\wedge s_2)^\beta}{(t-s_1\wedge s_2)^\beta} g(c(t-s_1\wedge s_2),z-x) \\
&\qquad \leq C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^\beta} g(c(t-s_1\vee s_2),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^\beta} g(c(t-s_1\wedge s_2),z-x)\right\}.
\end{aligned}$$

Combining the two previous cases, we get

$$\forall \beta \in [0,1], \quad |p_m(\mu,s_1,t,x,z) - p_m(\mu,s_2,t,x,z)| \\ \leq C\left\{ \frac{|s_1-s_2|^\beta}{(t-s_1)^\beta} g(c(t-s_1), z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^\beta} g(c(t-s_2), z-x) \right\}. \tag{4.37}$$

More generally, similar arguments as those previously employed allow to derive for $n = 0,\ 1,\ 2$

$$\forall \beta \in [0,1], \quad |\partial_x^n p_m(\mu,s_1,t,x,z) - \partial_x^n p_m(\mu,s_2,t,x,z)| \\ \leq C\left\{ \frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{n}{2}+\beta}} g(c(t-s_1), z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{n}{2}+\beta}} g(c(t-s_2), z-x) \right\}. \tag{4.38}$$

We now make use of the following decompositon

$$\partial_v^n[\partial_\mu p_{m+1}(\mu, s_1\vee s_2, t, x, z)](v) - \partial_v^n[\partial_\mu p_{m+1}(\mu, s_1\wedge s_2, t, x, z)](v) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}$$

with

$$\begin{aligned}
\mathrm{I} :=& \left\{ Df_{z-x}\left(\int_{s_1\vee s_2}^t a(r, [\bar X_r^{s_1\vee s_2,\xi,(m)}])dr \right) - Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r, [\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr \right) \right\} \\
&\cdot \int_{s_1\vee s_2}^t \left\{ \int (\widetilde a(r,y', [\bar X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde a(r,v,[\bar X_r^{s_1\vee s_2,\xi,(m)}]))\partial_x^{1+n} p_m(\mu, s_1\vee s_2, r, v, y')\, dy' \right. \\
&\quad \left. + \int\int (\widetilde a(r,y', [\bar X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde a(r,x', [\bar X_r^{s_1\vee s_2,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu, s_1\vee s_2, r, x', y')](v)\, dy'\, \mu(dx') \right\} dr, \\
\mathrm{II} :=& Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r, [\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr \right) \\
&\cdot \Big(\int_{s_1\vee s_2}^t \Big(\int (\widetilde a(r,y', [\bar X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde a(r,y',[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\, \partial_x^{1+n} p_m(\mu, s_1\vee s_2, r, v, y')\, dy' \\
&\quad + \int (\widetilde a(r,y', [\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,v,[\bar X_r^{s_1\wedge s_2,\xi,(m)}])) \\
&\qquad \cdot \left(\partial_x^{1+n} p_m(\mu, s_1\vee s_2, r, v, y') - \partial_x^{1+n} p_m(\mu, s_1\wedge s_2, r, v, y')\right)\, dy' \Big)\, dr \Big), \\
\mathrm{III} :=& Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r, [\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr \right) \\
&\cdot \Big(\int_{s_1\vee s_2}^t \left\{ \int (\widetilde a(r,y', [\bar X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde a(r,y',[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\, \partial_v^n[\partial_\mu p_m(\mu, s_1\vee s_2, r, x, y')](v)\, dy' \right. \\
&\quad \left. + \int\int (\widetilde a(r,y', [\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,x',[\bar X_r^{s_1\wedge s_2,\xi,(m)}])) \right. \\
&\qquad \left. \cdot (\partial_v^n[\partial_\mu p_m(\mu, s_1\vee s_2, r, x', y')](v) - \partial_v^n[\partial_\mu p_m(\mu, s_1\wedge s_2, r, x', y')](v))\, dy'\, \mu(dx') \right\}\, dr \Big), \\
\mathrm{IV} :=& Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r, [\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr \right) \\
&\cdot \Big(\int_{s_1\vee s_2}^t \left\{ \int (\widetilde a(r,y', [\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,v,[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_x^{1+n} p_m(\mu, s_1\wedge s_2, r, v, y')\, dy' \right. \\
&\quad \left. + \int\int (\widetilde a(r,y', [\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,x',[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu, s_1\wedge s_2, r, x', y')](v)\, dy'\, \mu(dx') \right\} dr \\
&- \int_{s_1\wedge s_2}^t \left\{ \int (\widetilde a(r,y', [\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,v,[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_x^{1+n} p_m(\mu, s_1\wedge s_2, r, v, y')\, dy' \right. \\
&\quad \left. + \int\int (\widetilde a(r,y', [\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,x',[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu, s_1\wedge s_2, r, x', y')](v)\, dy'\, \mu(dx') \right\} dr \Big)
\end{aligned}$$

and prove the following estimates

$$
\begin{aligned}
|\mathrm{I}| &\leq C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}}\, g(c(t-s_1),z-x)+\frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}}\, g(c(t-s_2),z-x)\right\}, \text{for all } \beta\in[0,1],\\
|\mathrm{II}| &\leq C\frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{\frac{1+n-\eta}{2}+\beta}}\, g(c(t-s_1\wedge s_2),z-x), \text{where } \beta\in[0,\frac{1+\eta}{2}), \text{for } n=0, \text{ and } \beta\in[0,\frac{\eta}{2}) \text{ for } n=1,\\
|\mathrm{III}| &\leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{\frac{1+n}{2}+\beta-\eta}}+\frac{1}{t-s_1\wedge s_2}\Big(\int_{s_1\vee s_2}^t\int\int(|y'-x'|^\eta\wedge 1)\\
&\quad\times|\partial_v^n[\partial_\mu p_m(\mu,s_1\vee s_2,r,x',y')](v)-\partial_v^n[\partial_\mu p_m(\mu,s_1\wedge s_2,r,x',y')](v)|\,dy'\,\mu(dx')\,dr\Big)\Big]\\
&\quad\times g(c(t-s_1\wedge s_2),z-x),\ \text{where } \beta\in[0,\frac{1}{2}+\eta), \text{if } n=0, \text{and } \beta\in[0,\eta) \text{if } n=1,\\
|\mathrm{IV}| &\leq C\frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{\frac{1+n-\eta}{2}+\beta}}\, g(c(t-s_1\wedge s_2),z-x) \text{ where } \beta\in[0,\frac{1+\eta}{2}), \text{for } n=0, \text{ and } \beta\in[0,\frac{\eta}{2}) \text{ for } n=1.
\end{aligned}
$$

The previous estimates in turn imply

$$
\begin{aligned}
&\int\int(|y'-x'|^\eta\wedge 1)|\partial_v^n[\partial_\mu p_{m+1}(\mu,s_1\vee s_2,t,x',y')](v)-\partial_v^n[\partial_\mu p_{m+1}(\mu,s_1\wedge s_2,t,x',y')](v)|\,dy'\,\mu(dx')\\
&\leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}+\frac{1}{(t-s_1\vee s_2)^{1-\frac{\eta}{2}}}\\
&\times\int_{s_1\vee s_2}^t\int\int(|y'-x'|^\eta\wedge 1)|\partial_v^n[\partial_\mu p_m(\mu,s_1\vee s_2,r,x',y')](v)-\partial_v^n[\partial_\mu p_m(\mu,s_1\wedge s_2,r,x',y')](v)|\,dy'\,\mu(dx')\,dr\Big]
\end{aligned}
$$

which, by an induction argument that we omit, yields

$$
\begin{aligned}
&\int\int(|z-x'|^\eta\wedge 1)|\partial_v^n[\partial_\mu p_{m+1}(\mu,s_1\vee s_2,t,x',z)](v)-\partial_v^n[\partial_\mu p_{m+1}(\mu,s_1\wedge s_2,t,x',z)](v)|\,dz\,\mu(dx')\\
&\qquad\qquad\leq C\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}
\end{aligned}
$$

for some positive constant $C:=C(T,\lambda,|\widetilde{a}|_\infty,|\widetilde{A}|_\infty,[\widetilde{a}]_H,[\widetilde{A}]_H)$. The previous bound finally implies (4.19). It now remains to obtain the announced estimates on I, II, III, IV. We only prove the estimates on I and II and omit the proof for the others since they stem from similar arguments.

∘ *Estimate on* I*:*

Using the fact that $y\mapsto\widetilde{a}(t,y,\mu)$ is η-Hölder uniformly with respect to t, μ, the space-time inequality (1.4) and (4.15), we get

$$
\begin{aligned}
&\Big|\int_{s_1\vee s_2}^t\Big\{\int(\widetilde{a}(r,y',[\bar{X}_r^{s_1\vee s_2,\xi,(m)}])-\widetilde{a}(r,v,[\bar{X}_r^{s_1\vee s_2,\xi,(m)}]))\partial_x^{1+n}p_m(\mu,s_1\vee s_2,r,v,y')\,dy'\\
&\quad+\int\int(\widetilde{a}(r,y',[\bar{X}_r^{s_1\vee s_2,\xi,(m)}])-\widetilde{a}(r,x',[\bar{X}_r^{s_1\vee s_2,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu,s_1\vee s_2,r,x',y')](v)\,dy'\,\mu(dx')\Big\}\,dr\Big|\\
&\leq C\int_{s_1\vee s_2}^t\frac{1}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}}}\,dr\\
&\leq C(t-s_1\vee s_2)^{1-\frac{(1+n-\eta)}{2}}.
\end{aligned}
$$

If $|s_1 - s_2| \geq t - s_1 \vee s_2$, one directly gets

$$\begin{aligned}\forall \beta \in [0,1], \quad &\Big|Df_{z-x}\Big(\int_{s_1\vee s_2}^t a(r,[\bar X_r^{s_1\vee s_2,\xi,(m)}])\,dr\Big)\Big| + \Big|Df_{z-x}\Big(\int_{s_1\wedge s_2}^t a(r,[\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr\Big)\Big| \\ &\leq C\Big[\frac{1}{(t-s_1\vee s_2)}g(c(t-s_1\vee s_2),z-x) + \frac{1}{(t-s_1\wedge s_2)}g(c(t-s_1\wedge s_2),z-x)\Big] \\ &\leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{1+\beta}}g(c(t-s_1\vee s_2),z-x) + \frac{(t-s_1\wedge s_2)^\beta}{(t-s_1\wedge s_2)^{1+\beta}}g(c(t-s_1\wedge s_2),z-x)\Big] \\ &\leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{1+\beta}}g(c(t-s_1\vee s_2),z-x) \\ &\qquad + \frac{(t-s_1\vee s_2)^\beta + |s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{1+\beta}}g(c(t-s_1\wedge s_2),z-x)\Big] \\ &\leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1)^{1+\beta}}g(c(t-s_1),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{1+\beta}}g(c(t-s_2),z-x)\Big].\end{aligned}$$

Otherwise if $|s_1 - s_2| \leq t - s_1 \vee s_2$, using the mean-value theorem, (**HR**) and (4.34), we obtain

$$\begin{aligned}&\Big|Df_{z-x}\Big(\int_{s_1\vee s_2}^t a(r,[\bar X_r^{s_1\vee s_2,\xi,(m)}])dr\Big) - Df_{z-x}\Big(\int_{s_1\wedge s_2}^t a(r,[\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr\Big)\Big| \\ &\leq C\int_0^1 \Big[\frac{|s_1-s_2|}{(t-(\lambda s_1+(1-\lambda)s_2)^2} + \frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{1+\beta-\frac{\eta}{2}}}\Big]\, g(c(t-(\lambda s_1+(1-\lambda)s_2),z-x)\,d\lambda \\ &\leq C\int_0^1 \Big[\frac{|s_1-s_2|^\beta}{(t-(\lambda s_1+(1-\lambda)s_2)^{1+\beta}} + \frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{1+\beta-\frac{\eta}{2}}}\Big]\, g(c(t-(\lambda s_1+(1-\lambda)s_2),z-x)\,d\lambda \\ &\leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{1+\beta}} + \frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{1+\beta-\frac{\eta}{2}}}\Big]\, g(c(t-s_1\wedge s_2),z-x).\end{aligned}$$

where we used the bound $(t-s)^{-1} \leq 2(t-s_1\wedge s_2)^{-1}$ for any $s \in (s_1,s_2)$ for the last inequality.

Hence, for all $\beta \in [0,1]$, one gets

$$\begin{aligned}&\Big|Df_{z-x}\Big(\int_{s_1\vee s_2}^t a(r,[\bar X_r^{s_1\vee s_2,\xi,(m)}])dr\Big) - Df_{z-x}\Big(\int_{s_1\wedge s_2}^t a(r,[\bar X_r^{s_1\wedge s_2,\xi,(m)}])dr\Big)\Big| \\ &\qquad\qquad \leq C\Big[\frac{|s_1-s_2|^\beta}{(t-s_1)^{1+\beta}}g(c(t-s_1),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{1+\beta}}g(c(t-s_2),z-x)\Big] \qquad (4.39)\end{aligned}$$

and gathering the previous computations we obtain the announced estimate on I.

∘ *Estimate on* II*:*

Using the fact that $\mu \mapsto a(t,\mu)$ belongs to ($\mathbf{CS}_+$), more precisely combining condition (2.9) with (4.35), one obtains

$$\begin{aligned}&\Big|\int_{s_1\vee s_2}^t \int (\widetilde a(r,y,[\bar X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde a(r,y,[\bar X_r^{s_1\wedge s_2,\xi,(m)}]))\,\partial_x^{1+n}p_m(\mu,s_1\vee s_2,r,v,y)\,dy\,dr\Big| \\ &\qquad\qquad \leq C|s_1-s_2|^\beta \int_{s_1\vee s_2}^t \frac{1}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\,dr \\ &\qquad\qquad \leq C\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1-n-\eta}{2}+\beta}}.\end{aligned}$$

Using the fact that $x \mapsto \widetilde a(r,x,\mu)$ is η-Hölder with (4.35) and the space-time inequality (1.4), one gets

$$\begin{aligned}&\Big|\int \Big(\widetilde a(r,y,[\bar X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde a(r,v,[\bar X_r^{s_1\wedge s_2,\xi,(m)}])\Big) \\ &\qquad .(\partial_x^{1+n}p_m(\mu,s_1\vee s_2,r,v,y) - \partial_x^{1+n}p_m(\mu,s_1\wedge s_2,r,v,y))\,dy\Big)\,dr\Big| \\ &\leq C\int (|y-v|^\eta \wedge 1)\,|\partial_x^{1+n}p_m(\mu,s_1\vee s_2,r,v,y) - \partial_x^{1+n}p_m(\mu,s_1\wedge s_2,r,v,y)|\,dy\,dr \\ &\leq C\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}.\end{aligned}$$

Gathering the two previous computations yield

$$|\mathrm{II}| \leq C\frac{|s_1-s_2|^\beta (t-s_1\vee s_2)^{1-\frac{(1+n-\eta)}{2}-\beta}}{t-s_1\wedge s_2}\, g(c(t-s_1\wedge s_2),z-x) \leq C\frac{|s_1-s_2|^\beta}{(t-s_1\wedge s_2)^{\frac{1+n-\eta}{2}+\beta}}\, g(c(t-s_1\wedge s_2),z-x)$$

which is the announced estimate on II.

Step 2: Extraction of a convergent subsequence.

Our next step now is to extract from the following sequences $\big\{\mathbb{L}^2 \ni \xi \mapsto \tilde{p}_m(\xi,s,t,x,z), m\geq 0\big\}$, $\big\{\mathbb{R}^d \ni v \mapsto \partial_\mu p_m(\mu,s,t,x,z)(v), m\geq 0\big\}$, $\big\{\mathbb{R}^d \ni v \mapsto \partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v), m\geq 0\big\}$ the corresponding subsequences which converge locally uniformly using the Arzelà-Ascoli theorem.

We remind the reader that since $(\mathbb{P}^{(m)})_{m\geq 0}$ converges weakly to $\mathbb{P}$, for any fixed $t>0$ and $z\in\mathbb{R}^d$, the sequence of functions $\{\mathcal{K}\ni(s,x,\mu)\mapsto p_m(\mu,s,t,x,z),\ m\geq 1\}$, $\mathcal{K}$ being a compact set of $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, converges to $(s,x,\mu)\mapsto\bar{p}(\mu,s,t,x,z)$, for any fixed (s,x,μ). Moreover, it is clearly uniformly bounded and from (4.30), (4.37) and the direct bound $\partial_x p_m(\mu,s,t,x,z)\leq C(t-s)^{-\frac{1}{2}}g(c(t-s),z-x)$, C, c being two positive constants independent of m, it is equicontinuous. Relabelling the indices if necessary, from the Arzelà-Ascoli theorem, we may assert that it converges uniformly. Hence, $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\bar{p}(\mu,s,t,x,z)$ is continuous.

For any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ and any integer m, the mapping $(s,x)\mapsto p_m(\mu,s,t,x,z)$ is in $\mathcal{C}^{0,2}([0,t)\times\mathbb{R}^d)$. Moreover, from the estimates (4.31), (4.38) and the bound $|\partial_x^3 p_m(\mu,s,t,x,z)|\leq C(t-s)^{-\frac{3}{2}}g(c(t-s),z-x)$, the sequence of functions $\mathcal{K}\ni(s,x,\mu)\mapsto\partial_x p_m(\mu,s,t,x,z)$, $\partial_x^2 p_m(\mu,s,t,x,z)$, are uniformly bounded and equicontinuous. Hence, from Arzelà-Ascoli's theorem, we may assert that $(s,x)\mapsto\bar{p}(\mu,s,t,x,z)\in\mathcal{C}^{0,2}([0,t)\times\mathbb{R}^d)$ and that the mappings $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\partial_x\bar{p}(\mu,s,t,x,z)$, $\partial_x^2\bar{p}(\mu,s,t,x,z)$ are continuous.

From (4.15), the sequence $\big\{\mathbb{L}^2\supset B(0,R)\ni\xi\mapsto D\tilde{p}_m(\xi,s,t,x,z)=\partial_\mu p_m(\mu,s,t,x,z)(\xi),\ m\geq 1\big\}$ is uniformly bounded and equicontinuous for any $R>0$. Relabelling the indices if necessary, from the Arzelà-Ascoli theorem, we may assert it converges uniformly. Hence, $L^2(\Omega,\mathcal{A},\mathbb{P})\supset B(0,R)\ni\xi\mapsto\tilde{p}(\xi,s,t,x,z)$ is differentiable. Passing to the limit (along the considered subsequence) in (4.22) or (4.23) as $m\uparrow\infty$, we obtain that $(\mu,v)\mapsto\partial_\mu\bar{p}(\mu,s,t,x,z)(v)$ satisfies

$$\begin{aligned}\partial_\mu\bar{p}(\mu,s,t,x,z)(v) &:= Df_{z-x}\left(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr\right)\\ &\quad\cdot\int_s^t\Big\{\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi}])-\widetilde{a}(r,v,[\bar{X}_r^{s,\xi}]))\partial_x\bar{p}(\mu,s,r,v,y')\,dy'\\ &\quad+\int\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi}])-\widetilde{a}(r,x',[\bar{X}_r^{s,\xi}]))\partial_\mu\bar{p}(\mu,s,r,x',y')(v)\,dy'\,\mu(dx')\Big\}\,dr\end{aligned}\tag{4.40}$$

and the estimate (4.7) holds for $n=0$. As a consequence, $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto\bar{p}(\mu,s,t,x,z)$ is continuously L-differentiable and its derivative satisfies (4.40). From the estimates (4.19), (4.18) (both for $n=0$) and (4.15), the same conclusion holds for the sequence $\mathcal{K}\ni(s,x,\mu,v)\mapsto\partial_\mu p_m(\mu,s,t,x,z)(v)$, $m\geq 1$, $\mathcal{K}$ being a compact set of $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$, that is, it is uniformly bounded and equicontinuous (equicontinuity w.r.t the space variable x being a direct consequence of (4.23)) so that the map $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(s,x,\mu,v)\mapsto\partial_\mu\bar{p}(\mu,s,t,x,z)(v)$ is continuous. By passing to the limit in (4.18) ($n=0$), we obtain (4.10) ($n=0$).

From the estimates (4.15) and (4.17), the sequence $\mathbb{R}^d\supset B(0,R)\ni v\mapsto\partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v)$, $m\geq 1$, is bounded and equicontinuous. From the Arzelà-Ascoli theorem, we thus conclude that $v\mapsto\partial_\mu\bar{p}(\mu,s,t,x,z)(v)$ is continuously differentiable. Passing to the limit (along the considered subsequence) in (4.24), we get

$$\begin{aligned}&\partial_v[\partial_\mu\bar{p}(\mu,s,t,x,z)](v)\\ &= Df_{z-x}(\int_s^t a(r,[\bar{X}_r^{s,\xi}])dr).\int_s^t\Big\{\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi}])-\widetilde{a}(r,v,[\bar{X}_r^{s,\xi}]))\partial_x^2\bar{p}(\mu,s,r,v,y')\,dy'\\ &+\int\int(\widetilde{a}(r,y',[\bar{X}_r^{s,\xi}])-\widetilde{a}(r,x',[\bar{X}_r^{s,\xi}]))\partial_v[\partial_\mu\bar{p}(\mu,s,r,x',y')](v)dy'\,\mu(dx')\Big\}\,dr.\end{aligned}\tag{4.41}$$

and the estimates (4.7) for $n=1$ and (4.9) hold. The continuity of the map $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(s,x,\mu,v)\mapsto\partial_v[\partial_\mu\bar{p}(\mu,s,t,x,z)](v)$ can be deduced from the uniform convergence of the sequence of continuous mappings $\mathcal{K}\ni(s,x,\mu,v)\mapsto\partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v)$, $m\geq 1$, $\mathcal{K}$ being a compact set of $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$, along a subsequence, obtained from the estimates (4.17), (4.18) and (4.19) for

$n = 1$ (the equicontinuity w.r.t the space variable being a direct consequence of (4.24)), combined with the Arzelà-Ascoli theorem. The estimates (4.10) for $n = 1$, (4.11) follow by passing to the limit in the corresponding upper-bounds proved in the first step.

Step 3: $\mathcal{C}^{1,2,2}([0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ regularity and related time estimates.

Let us now prove that $(s,x,\mu) \mapsto \bar{p}(\mu,s,t,x,z)$ is in $\mathcal{C}^{1,2,2}([0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$. From the Markov property satisfied by the SDE (4.1), stemming from the well-posedness of the related martingale problem, the following relation is satisfied for all $h > 0$

$$\bar{p}(\mu, s-h, t, x, z) = \mathbb{E}[\bar{p}([X_s^{s-h,\xi}], s, t, X_s^{s-h,x,\mu}, z)].$$

From the very definition of the map $x \mapsto \bar{p}(\mu,s,t,x,z)$, one has $|\partial_x \bar{p}(\mu,s,t,x,z)| \leq C(t-s)^{-1/2} g(c(t-s), z-x)$ so that, combining estimates (4.7) with the chain rule formula of Proposition 2.1 (with respect to the space and measure variables only) we obtain

$$\mathbb{E}[\bar{p}([X_s^{s-h,\xi}], s, t, X_s^{s-h,x,\mu}, z)] = \bar{p}(\mu,s,t,x,z) + \mathbb{E}\left[\int_{s-h}^{s} \bar{\mathcal{L}}_r \bar{p}([\bar{X}_r^{s-h,\xi}], s, t, \bar{X}_r^{s-h,x,\mu}, z)\, dr\right]$$

with $\bar{\mathcal{L}}_r g(x,\mu) := \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(r,\mu)\partial_{x_i,x_j} g(x,\mu) + \frac{1}{2}\int \sum_{i,j=1}^d a_{i,j}(r,\mu)\partial_{v_i}[\partial_\mu g(x,\mu)(v)]_j\, \mu(dv)$. Hence, one has

$$\frac{1}{h}(\bar{p}(\mu,s-h,t,x,z) - \bar{p}(\mu,s,t,x,z)) = \frac{1}{h}\mathbb{E}\left[\int_{s-h}^{s} \bar{\mathcal{L}}_r \bar{p}([\bar{X}_r^{s-h,\xi}], s, t, \bar{X}_r^{s-h,x,\mu}, z)\, dr\right]$$

with

$$\begin{aligned}\bar{\mathcal{L}}_r \bar{p}([\bar{X}_r^{s-h,\xi}], s, t, x, z) &= \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(r,[\bar{X}_r^{s-h,\xi}])\partial_{x_i,x_j}\bar{p}([\bar{X}_r^{s-h,\mu}], s, t, x, z) \\ &+ \frac{1}{2}\int\int \sum_{i,j=1}^d a_{i,j}(r,[\bar{X}_r^{s-h,\xi}])\partial_{v_i}[\partial_\mu \bar{p}([\bar{X}_r^{s-h,\mu}], s, t, x, z)(v)]_j\, \bar{p}(\mu, s-h, r, x', v)\, dv\, \mu(dx').\end{aligned}$$

Letting $h \downarrow 0$, from the boundedness and continuity of the coefficients as well as the continuity of the maps $(\mu,x) \mapsto \bar{p}(\mu,s,t,x,z)$, $\partial_x^2 \bar{p}(\mu,s,t,x,z)$, $\partial_v[\partial_\mu \bar{p}(\mu,s,t,x,z)]$, we deduce that $[0,t) \ni s \mapsto \bar{p}(\mu,s,t,x,z)$ is left differentiable. Still from the continuity of the coefficients and of the map $(s,x,\mu) \mapsto \bar{\mathcal{L}}_s \bar{p}(\mu,s,t,x,z)$, we then conclude that it is differentiable in time on the interval $[0,t)$ with a time derivative satisfying

$$\partial_s \bar{p}(\mu,s,t,x,z) = -\bar{\mathcal{L}}_s \bar{p}(\mu,s,t,x,z) \quad \text{ on } [0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d).$$

The time derivative estimates (4.8) and (4.12) now follow from the previous relation (4.7), the inequality $|\partial_x^n \bar{p}(\mu,s,t,x,z)| \leq C(t-s)^{-\frac{n}{2}} g(c(t-s), z-x)$, $n = 1,2$ as well the estimates (4.11) and (4.38).

□

4.3. **Well-posedness of the martingale problem.** We now have all the ingredients to prove the well-posedness of the original martingale problem. Under (**HC**), a standard compactness argument implies weak existence of solutions to the SDE (1.1). For sake of completeness, we provide a simple proof of this claim in Appendix, Section 7.1. We consider two probability measures $\mathbb{P}^1$ and $\mathbb{P}^2$ on the space $(\mathcal{C}([0,\infty),\mathbb{R}^d), \mathcal{B}(\mathcal{C}([0,\infty),\mathbb{R}^d))$ induced by two weak solutions to the SDE (1.1). The time marginals at time s are denoted by $\mathbb{P}^1(s)$ and $\mathbb{P}^2(s)$ respectively. We also introduce the two probability measures $\widetilde{\mathbb{P}}^i_{s,x}$, $i = 1, 2$, induced by the two unique weak solutions to the SDE with dynamics

$$X_t^{s,x,\mu} = x + \int_s^t b(r, X_r^{s,x,\mu}, \mathbb{P}^i(r))dr + \int_s^t \sigma(r, X_r^{s,x,\mu}, \mathbb{P}^i(r))dW_r, \quad i = 1,2.$$

Note that weak uniqueness to the above SDE under (**HR**) and (**HE**) follows from well-known results, see e.g. Stroock and Varadhan [SV79]. Importantly, we remark that for any $t \geq 0$, $\mathbb{P}^i(t)$ is the pushforward measure of μ by the map $x \mapsto \widetilde{\mathbb{P}}^i_{s,x}(t)$, that is, $\mathbb{P}^i(t) = \int \widetilde{\mathbb{P}}^i_{s,x}(t)\, \mu(dx)$, or in other words, $\mathbb{E}_{\mathbb{P}^i}[h(y_t)] = \int \mathbb{E}_{\widetilde{\mathbb{P}}^i_{s,x}}[h(y_t)]\, \mu(dx)$ for any bounded measurable functions h. In what follows, to make the notation simpler, we will simply write $\widetilde{\mathbb{P}}^i$ instead of $\widetilde{\mathbb{P}}^i_{s,x}$ when there is no ambiguity.

We consider the resolvents associated to the above SDEs defined for all bounded measurable $h : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ and all $\lambda > 0$ by

$$S^i_\lambda h(s,x,\mu) := \mathbb{E}_{\widetilde{\mathbb{P}^i}}\Big[\int_s^\infty e^{-\lambda(t-s)} h(t,y_t,\mathbb{P}^i(t))\,dt\Big] = \int_s^\infty e^{-\lambda(t-s)}\mathbb{E}_{\widetilde{\mathbb{P}^i}}\Big[h(t,y_t,\mathbb{P}^i(t))\Big]\,dt, \quad i=1,2,$$
$$S^\Delta_\lambda h(s,x,\mu) := (S^1_\lambda - S^2_\lambda)h(s,x,\mu),$$
$$\|S^\Delta_\lambda\| := \sup_{|h|_\infty \le 1} |S^\Delta_\lambda h|$$

where we emphasise that the supremum in the last definition is taken over all bounded continuous function h that *does not depend on the time and the measure variables.*

Clearly, one has $\|S^\Delta_\lambda\| \le 2/\lambda < \infty$ for $\lambda > 0$. For a given $y \in \mathbb{R}^d$, we also define the two operators

$$\mathcal{L}_t h(t,x,\mu) = \sum_{i=1}^d b_i(t,x,\mu)\partial_{x_i} h(t,x,\mu) + \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,x,\mu)\partial_{x_i,x_j} h(t,x,\mu)$$
$$+ \int \left\{ \sum_{i=1}^d b_i(t,z,\mu)[\partial_\mu h(t,x,\mu)(z)]_i + \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,z,\mu)\partial_{z_i}[\partial_\mu h(t,x,\mu)(z)]_j \right\}\mu(dz),$$
$$\bar{\mathcal{L}}_t h(t,x,\mu) = \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,y,\mu)\partial_{x_i,x_j} h(t,x,\mu) + \int \left\{ \frac{1}{2}\sum_{i,j=1}^d a_{i,j}(t,y,\mu)\partial_{z_i}[\partial_\mu h(t,x,\mu)(z)]_j \right\}\mu(dz)$$

both acting on smooth test functions $h \in \mathcal{C}^{1,2,2}(\mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$. Importantly, note that $\bar{\mathcal{L}}_t$ has been obtained from $\mathcal{L}_t$ by removing the drift and freezing the diffusion coefficient in the space variable but *not in the measure argument.*

From the chain rule formula of Proposition 5.102 of [CD18], applied to $h(t,y_t,\mathbb{P}^i(t))$ under $\widetilde{\mathbb{P}}^i$ for $h \in \mathcal{C}^{1,2,2}(\mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)))$, h having bounded continuous derivatives and the coefficients being bounded and continuous, the following identities hold

$$\begin{aligned} S^i_\lambda(\lambda - (\partial_t + \mathcal{L}))h(s,x,\mu) &= \int_s^\infty e^{-\lambda(t-s)}\mathbb{E}_{\widetilde{\mathbb{P}^i}}[(\lambda - (\partial_t + \mathcal{L}_t))h(t,y_t,\mathbb{P}^i(t))]\,dt \\ &= \int_s^\infty e^{-\lambda(t-s)}(\lambda\mathbb{E}_{\widetilde{\mathbb{P}^i}}[h(t,y_t,\mathbb{P}^i(t))] - \partial_t\mathbb{E}_{\widetilde{\mathbb{P}^i}}[h(t,y_t,\mathbb{P}^i(t))])\,dt \\ &= h(s,x,\mu), \quad i=1,2. \end{aligned} \tag{4.42}$$

In a similar fashion, we define the resolvent of the process with frozen diffusion coefficient (with respect to the space variable) at $y \in \mathbb{R}^d$ by

$$\bar{R}_\lambda h(s,x,\mu) = \int_s^\infty e^{-\lambda(t-s)}\mathbb{E}[h(t,\bar{X}^{s,x,\mu}_t,[\bar{X}^{s,\xi}_t])]\,dt$$

where $(\bar{X}^{s,x,\mu}_t, \bar{X}^{s,\xi}_t)_{t\ge s}$ is the approximation process defined as the unique weak solution (see Lemma 4.1) to the SDE with dynamics

$$\bar{X}^{s,\xi}_t = \xi + \int_s^t \sigma(r,y,[\bar{X}^{s,\xi}_r])\,dW_r, \tag{4.43}$$

$$\bar{X}^{s,x,\mu}_t = x + \int_s^t \sigma(r,y,[\bar{X}^{s,\xi}_r])\,dW_r \tag{4.44}$$

as well as the Markov semigroup operator $\bar{P}_{s,t}h(x,\mu) = \mathbb{E}[h(\bar{X}^{s,x,\mu}_t,[\bar{X}^{s,\xi}_t])]$ acting on bounded measurable functions h defined on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$.

Note that, to simplify the notation, in what follows we omit to specify the dependence with respect to the point y in the dynamics of the process $(\bar{X}^{s,x,\mu}_t, \bar{X}^{s,\xi}_t)_{t\ge s}$ as well as in the semigroup operator $\bar{P}_{s,t}$ when there is no ambiguity. The chain rule formula applied to $h(t,\bar{X}^{s,x,\mu}_t,[\bar{X}^{s,\xi}_t])$, similarly to (4.42), implies

$$\bar{R}_\lambda(\lambda - (\partial_t + \bar{\mathcal{L}}))h(s,x,\mu) = h(s,x,\mu).$$

The Markov property satisfied by the solution to the SDE (4.43) yields $[\bar{X}^{s,\bar{X}^{r,\xi}_s}_t] = [\bar{X}^{r,\xi}_t]$, for any $r \le s \le t$. As a consequence, from the very definition of the dynamics (4.44), we deduce that $[\bar{X}^{s,\bar{X}^{r,x,[\xi]}_s,[\bar{X}^{r,\xi}_s]}_t] = [\bar{X}^{r,x,\mu}_t]$.

For a fixed $t > 0$, we introduce the real-valued map $[0,t] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (s,x,\mu) \mapsto V(s,x,\mu) := \mathbb{E}[h(\bar{X}_t^{s,x,\mu}, [\bar{X}_t^{s,\xi}])]$. The preceding discussion clearly yields

$$V(s, \bar{X}_s^{r,x,\mu}, [\bar{X}_s^{r,\xi}]) = \mathbb{E}[h(\bar{X}_t^{s,\bar{X}_s^{r,x,\mu},[\bar{X}_s^{r,\xi}]}, [\bar{X}_t^{s,\bar{X}_s^{r,\xi}}])|\mathcal{F}_s] = \mathbb{E}[h(\bar{X}_t^{r,x,\mu}, [\bar{X}_t^{r,\xi}])|\mathcal{F}_s]$$

is a continuous martingale.

From now on, we restrict our consideration to the class of bounded continuous test functions h that do not depend on the time and the measure variables, that is, $h(t,x,\mu) = h(x)$. Hence, one has $V(s,x,\mu) = \int_{\mathbb{R}^d} h(z)\bar{p}(\mu,s,t,x,z)\,dz$. By continuity and boundedness of $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \ni (s,x,\mu) \mapsto \bar{p}(\mu,s,t,x,z)$ and h, the map $[0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (s,x,\mu) \mapsto V(s,x,\mu)$ is continuous and satisfies $\lim_{s\uparrow t} V(s,x,\mu) = h(x)$. We also remark that, from Proposition 4.1, $(s,x,\mu) \mapsto V(s,x,\mu) \in \mathcal{C}^{1,2,2}([0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$, so that, applying the chain rule formula (2.5) to $V(s, \bar{X}_s^{r,x,\mu}, [\bar{X}_s^{r,\xi}])$, $0 \le r \le s < t$, taking expectation, then differentiating with respect to s and finally letting $s \downarrow r$, we get that $(s,x,\mu) \mapsto V(s,x,\mu)$ satisfies the PDE:

$$\begin{cases} (\partial_s + \bar{\mathcal{L}}_s)V(s,x,\mu) = 0, \quad \text{for } (s,x,\mu) \in [0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d), \\ V(t,x,\mu) = h(x), \quad \text{for } (x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d). \end{cases}$$

From the previous argument and using the estimates of Proposition 4.1 to interchange differentiations and integrals, the following identities hold

$$\begin{aligned} (\lambda - (\partial_t + \bar{\mathcal{L}}))\bar{R}_\lambda h(s,x,\mu) &= \int_s^\infty e^{-\lambda(t-s)}\lambda\mathbb{E}[h(\bar{X}_t^{s,x,\mu})]\,dt - \partial_s \int_s^\infty e^{-\lambda(t-s)}\mathbb{E}[h(\bar{X}_t^{s,x,\mu})]\,dt \\ &\quad - \int_s^\infty e^{-\lambda(t-s)}\bar{\mathcal{L}}_s\mathbb{E}[h(\bar{X}_t^{s,x,\mu})]\,dt \\ &= \int_s^\infty e^{-\lambda(t-s)}\lambda\mathbb{E}[h(\bar{X}_t^{s,x,\mu})]\,dt + h(x) - \int_s^\infty e^{-\lambda(t-s)}\lambda\mathbb{E}[h(\bar{X}_t^{s,x,\mu})]\,dt \\ &\quad - \int_s^\infty e^{-\lambda(t-s)}(\partial_s + \bar{\mathcal{L}}_s)\mathbb{E}[h(\bar{X}_t^{s,x,\mu})]\,dt \\ &= h(x) \end{aligned} \tag{4.45}$$

for $h \in \mathcal{C}_b^2(\mathbb{R}^d)$. Let $\delta_\varepsilon^y(x) = g(\varepsilon, y-x)$, $\varepsilon > 0$, be an approximation of the Dirac mass at point y. Let $r > 0$. Since $(s,x,\mu) \mapsto \bar{p}(\mu,s,t+r,x,z) \in \mathcal{C}^{1,2,2}([0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$, we remark that for $h \in \mathcal{C}_b(\mathbb{R}^d)$, the mapping

$$\begin{aligned} (s,x,\mu) \mapsto \bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h(s,x,\mu) &= \int_s^\infty e^{-\lambda(t-s)}\mathbb{E}[\bar{P}_{t,t+r}\delta_\varepsilon^y h(\bar{X}_t^{s,x,\mu}, [\bar{X}_t^{s,\xi}])]\,dt \\ &= \int_s^\infty e^{-\lambda(t-s)}\bar{P}_{s,t+r}\delta_\varepsilon^y h(x,\mu)\,dt \end{aligned}$$

is in $\mathcal{C}^{1,2,2}(\mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$. Note that the fact that $r > 0$ allows one to differentiate w.r.t the variables s or x inside the integrals by removing the singularity when $t \downarrow s$. Consequently, from the identity (4.42) and Fubini's theorem, one derives

$$S_\lambda^i \int_{\mathbb{R}^d} (\lambda - (\partial_t + \mathcal{L}))\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy = \int_{\mathbb{R}^d} S_\lambda^i(\lambda - (\partial_t + \mathcal{L}))\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy = \int \bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy, \quad i = 1,2,$$

Combining the previous equality with the identity (4.45) and Fubini's theorem we thus obtain

$$\begin{aligned} \int \bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy &= \int_{\mathbb{R}^d} S_\lambda^i(\lambda - (\partial_t + \mathcal{L}))\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy \\ &= \int_{\mathbb{R}^d} S_\lambda^i(\lambda - (\partial_t + \bar{\mathcal{L}}))\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy - \int_{\mathbb{R}^d} S_\lambda^i(\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy \\ &= \int_{\mathbb{R}^d} S_\lambda^i \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy - \int_{\mathbb{R}^d} S_\lambda^i(\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy \\ &= S_\lambda^i \int_{\mathbb{R}^d} \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy - S_\lambda^i \int_{\mathbb{R}^d} (\mathcal{L} - \bar{\mathcal{L}})\bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy. \end{aligned} \tag{4.46}$$

We now pass to the limit in the previous equality as $\varepsilon \downarrow 0$ and then as $r \downarrow 0$. Let us consider the term appearing in the left-hand side of the previous expression. The dominated convergence and the

continuity of the map $r \mapsto \bar{p}^y(\mu, s, t+r, y)$ yield

$$\begin{aligned}\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\int \bar{R}_\lambda \bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h(s,x,\mu)\,dy &= \lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\int_s^\infty e^{-\lambda(t-s)}\int\int h(z)\bar{p}^y(\mu,s,t+r,x,z)\,g(\varepsilon,y-z)\,dz\,dy\\ &= \lim_{r\downarrow 0}\int_s^\infty e^{-\lambda(t-s)}\int\int h(y)\bar{p}^y(\mu,s,t+r,x,y)\,dy\\ &= \int_s^\infty e^{-\lambda(t-s)}\int h(y)\bar{p}^y(\mu,s,t,x,y)\,dy\\ &=: \widehat{R}_\lambda h(s,x,\mu).\end{aligned}$$

We then apply the dominated convergence theorem and Lemma 7.1 to obtain

$$\begin{aligned}\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\int \bar{P}_{t,t+r}\delta_\varepsilon^y h(x,\mu)dy &= \lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\int\int h(z)\bar{p}^y(\mu,t,t+r,y)\,g(\varepsilon,y-z)\,dz\,dy\\ &= \lim_{r\downarrow 0}\int h(y)\bar{p}^y(\mu,t,t+r,x,y)dy\\ &= h(x)\end{aligned}$$

so that, again by dominated convergence,

$$\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0} S_\lambda^i\left(\int \bar{P}^y_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy\right) = S_\lambda^i h. \tag{4.47}$$

In order to investigate the limit of the second term appearing in the right-hand side of (4.46), we first need to introduce some notations and derive some key estimates. Observe that from the very definition of the operators $\mathcal{L}_s$, $\bar{\mathcal{L}}_s$ and the density $z \mapsto \bar{p}(\mu, s, t, x, z)$, one can write

$$\begin{aligned}(\mathcal{L}_s - \bar{\mathcal{L}}_s)\bar{P}_{s,t}h(x,\mu) &= \int_{\mathbb{R}^d} h(z)\,(\mathcal{L}_s - \bar{\mathcal{L}}_s)\bar{p}(\mu,s,t,x,z)\,dz\\ &= \int_{\mathbb{R}^d} h(z)\,(\mathcal{S}^{1,y}_{s,t}(x,\mu,z) + \mathcal{S}^{2,y}_{s,t}(x,\mu,z))\,dz\end{aligned}$$

with

$$\begin{aligned}\mathcal{S}^{1,y}_{s,t}(x,\mu,z) &:= \Big(-\sum_{i=1}^d b_i(s,x,\mu)\,\partial_{x_i}\bar{p}(\mu,s,t,x,z)\\ &\quad + \frac{1}{2}\sum_{i,j=1}^d (a_{i,j}(s,x,\mu) - a_{i,j}(s,y,\mu))\,\partial^2_{x_i,x_j}\bar{p}(\mu,s,t,x,z)\\ &= \Big(-\sum_{i=1}^d b_i(s,x,\mu)H_1^i\left(\int_s^t a(r,y,[\bar{X}_r^{s,\xi}])dr, z-x\right)\\ &\quad + \frac{1}{2}\sum_{i,j=1}^d (a_{i,j}(s,x,\mu) - a_{i,j}(s,y,\mu))H_2^{i,j}\left(\int_s^t a(r,y,[\bar{X}_r^{s,\xi}])dr, z-x\right)\Big)\\ &\quad \times g\left(\int_s^t a(r,y,[\bar{X}_r^{s,\xi}])dr, z-x\right),\end{aligned} \tag{4.48}$$

$$
\begin{aligned}
\mathcal{S}^{2,y}_{s,t}(x,\mu,z) &:= \int \sum_{i=1}^{d} b_i(s,z',\mu)[\partial_\mu \bar{p}(\mu,s,t,x,z)(z')]_i \,\mu(dz') \\
&\quad + \int \frac{1}{2} \sum_{i,j=1}^{d} \left(a_{i,j}(s,z',\mu) - a_{i,j}(s,y,\mu)\right) \partial_{z'_j}[\partial_\mu \bar{p}(\mu,s,t,x,z)(z')]_i \,\mu(dz') \\
&= \int \left\{ \sum_{i,j,k=1}^{d} b_i(s,z',\mu) \frac{1}{2} (H_2^{j,k}.g)\left(\int_s^t a(r,y,[\bar{X}_r^{s,\xi}])dr, z-x\right) \right. \\
&\quad \times \int_s^t \left(\partial_\nu \Big[\int \widetilde{a}_{j,k}(r,y,z'',[\bar{X}_r^{s,\xi}])\, \bar{p}^y(\nu,s,r,z'')\, dz'' \Big]_{\nu=\mu}(z')\right)_i dr \qquad (4.49) \\
&\quad + \frac{1}{2} \sum_{i,j,k,\ell=1}^{d} \left(a_{i,j}(s,z',\mu) - a_{i,j}(s,y,\mu)\right) \frac{1}{2} (H_2^{k,\ell}.g)\left(\int_s^t a(r,y,[\bar{X}_r^{s,\xi}])dr, z-x\right) \\
&\quad \left. \times \int_s^t \partial_{z'_j} \left(\partial_\nu \left[\int \widetilde{a}_{k,\ell}(r,y,z'',[\bar{X}_r^{s,\xi}])\, \bar{p}^y(\nu,s,r,z'')\, dz'' \right]_{\nu=\mu}(z')\right)_i dr \right\} \mu(dz').
\end{aligned}
$$

When the freezing point y is chosen to be the terminal point z of the above kernels, we denote by $\mathcal{S}_{s,t}(x,\mu,z) = \mathcal{S}^{1,z}_{s,t}(x,\mu,z) + \mathcal{S}^{2,z}_{s,t}(x,\mu,z)$ the corresponding kernel and also write $\mathcal{S}^{1,y}_{s,t}h(x,\mu) := \int_{\mathbb{R}^d} h(z)\, \mathcal{S}^{1,y}_{s,t}(x,\mu,z)\, dz$, $\mathcal{S}^{2,y}_{s,t}h(x,\mu) = \int_{\mathbb{R}^d} h(z)\, \mathcal{S}^{2,y}_{s,t}(x,\mu,z)\, dz$ and $\mathcal{S}_{s,t}h(x,\mu) = \int_{\mathbb{R}^d} h(z)\, \mathcal{S}^{z}_{s,t}(x,\mu,z)\, dz$ for the associated operators.

Importantly, we point out that the law $[\bar{X}_r^{s,\xi}]$ also depends on the freezing point y in the above kernels. We now derive some important estimates on these kernels. First let us note that from the boundedness of the coefficients and (1.5) one easily gets that there exists positive constants $C := C(T, |b|_\infty, |a|_\infty)$, $c := c(\lambda)$ such that

$$
\forall (\mu,x,y,z) \in \mathcal{P}_2(\mathbb{R}^d) \times (\mathbb{R}^d)^3, \quad |\mathcal{S}^{1,y}_{s,t}(x,\mu,z)| \le C(t-s)^{-1} g(c(t-s), z-x) \tag{4.50}
$$

and, from assumption (**HR**) and the space-time inequality (1.4), one also derives

$$
\forall (\mu,x,z) \in \mathcal{P}_2(\mathbb{R}^d) \times (\mathbb{R}^d)^2, \quad |\mathcal{S}^{1,z}_{s,t}(x,\mu,z)| \le C(t-s)^{-1+\frac{\eta}{2}} g(c(t-s), z-x). \tag{4.51}
$$

The estimate on $\mathcal{S}^{2,y}_{s,t}(x,\mu,z)$ follows directly from Proposition 4.1. Using the boundedness of the coefficients, one easily gets

$$
\forall (\mu,x,y,z) \in \mathcal{P}_2(\mathbb{R}^d) \times (\mathbb{R}^d)^3, \quad |\mathcal{S}^{2,y}_{s,t}(x,\mu,z)| \le C(t-s)^{-1+\frac{\eta}{2}} g(c(t-s), z-x). \tag{4.52}
$$

Observe now that, by Fubini's theorem, one has

$$
\begin{aligned}
\int (\mathcal{L} - \bar{\mathcal{L}}) \bar{R}_\lambda \bar{P}^y_{\cdot,\cdot+r} \delta^y_\varepsilon h(x,\mu)\, dy &= \int_s^\infty \int e^{-\lambda(t-s)} (\mathcal{L}_s - \bar{\mathcal{L}}_s) \bar{P}_{s,t+r} \delta^y_\varepsilon h(x,\mu)\, dt\, dy \\
&= \int_s^\infty e^{-\lambda(t-s)} \int (\mathcal{S}^{1,y}_{s,t+r} + \mathcal{S}^{2,y}_{s,t+r}) \delta^y_\varepsilon h(x,\mu)\, dy\, dt \\
&= \int_s^\infty e^{-\lambda(t-s)} \int_{(\mathbb{R}^d)^2} \delta^y_\varepsilon(z) h(z)\, (\mathcal{S}^{1,y}_{s,t+r} + \mathcal{S}^{2,y}_{s,t+r})(x,\mu,z)\, dz\ dy\, dt.
\end{aligned}
$$

Combining (4.50), (4.52) with Lebesgue's dominated convergence theorem yields

$$
\begin{aligned}
\lim_{\varepsilon \downarrow 0} \int_{(\mathbb{R}^d)^2} (\mathcal{S}^{1,y}_{s,t+r} + \mathcal{S}^{2,y}_{s,t+r})(x,\mu,z) \delta^y_\varepsilon(z) h(z) dz\, dy &= \int_{\mathbb{R}^d} (\mathcal{S}^{1,y}_{s,t+r} + \mathcal{S}^{2,y}_{s,t+r})(x,\mu,y) h(y) dy \\
&= \mathcal{S}_{s,t+r} h(x,\mu)
\end{aligned}
$$

so that again by (4.51), (4.52) and the Lebesgue dominated convergence theorem

$$
\lim_{\varepsilon \downarrow 0} \int (\mathcal{L} - \bar{\mathcal{L}}) \bar{R}_\lambda \bar{P}_{\cdot,\cdot+r} \delta^y_\varepsilon h(x,\mu)\, dy = \int_s^\infty e^{-\lambda(t-s)} \mathcal{S}_{s,t+r} h(x,\mu)\, dt.
$$

We finally observe that $\mathbb{R}_+ \ni r \mapsto \mathcal{S}_{s,t+r}h(x,\mu)$ is continuous so that from the estimates (4.51), (4.52) and Lebesgue's dominated convergence theorem, we derive

$$\begin{aligned}\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}\int(\mathcal{L}-\bar{\mathcal{L}})\bar{R}_\lambda\bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h(x,\mu)\,dy &= \lim_{r\downarrow 0}\int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t+r}h(x,\mu)\,dt\\ &= \int_s^\infty e^{-\lambda(t-s)}\lim_{r\downarrow 0}\mathcal{S}_{s,t+r}h(x,\mu)\,dt\\ &= \int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t}h(x,\mu)\,dt\end{aligned}$$

which in turn, again by the dominated convergence, yields

$$\lim_{r\downarrow 0}\lim_{\varepsilon\downarrow 0}S_\lambda^i\left(\int(\mathcal{L}-\bar{\mathcal{L}})\bar{R}_\lambda\bar{P}_{\cdot,\cdot+r}\delta_\varepsilon^y h\,dy\right) = S_\lambda^i\left(\int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t}h\,dt\right).$$

Hence, putting everything together we conclude that the following key identity holds

$$S_\lambda^i h - S_\lambda^i\left(\int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t}h\,dt\right) = \widehat{R}_\lambda h(s,x,\mu),\ i=1,2 \tag{4.53}$$

which in turn implies

$$S_\lambda^\Delta h = S_\lambda^\Delta\left(\int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t}h\,dt\right)$$

and, from (4.51) and (4.52), one can pick λ large enough such that

$$\left|\int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t}h\,dt\right| \le |h|_\infty\int_s^\infty e^{-\lambda(t-s)}(t-s)^{-1+\frac{\eta}{2}}dt = |h|_\infty\frac{\Gamma(\frac{\eta}{2})}{\lambda^{\frac{\eta}{2}}} < \frac{1}{2}|h|_\infty$$

which by the very definition of $\|S_\lambda^\Delta\|$ clearly yields

$$\left|S_\lambda^\Delta h\right| = \left|S_\lambda^\Delta\left(\int_s^\infty e^{-\lambda(t-s)}\mathcal{S}_{s,t}h\,dt\right)\right| \le \frac{1}{2}\|S_\lambda^\Delta\||h|_\infty.$$

By an approximation argument, the last inequality remains valid for real-valued bounded continuous functions. Taking the supremum over h (we remind the reader that h does not depend on the time and the measure variables) satisfying $|h|_\infty \le 1$ yields $\|S_\lambda^\Delta\| \le \frac{1}{2}\|S_\lambda^\Delta\|$ and, since $\|S_\lambda^\Delta\| < \infty$, for any $\lambda > 0$, we conclude that $\|S_\lambda^\Delta\| = 0$. Consequently, $\int_s^\infty e^{-\lambda(t-s)}\mathbb{E}_{\widetilde{\mathbb{P}}^1}[h(y_t)]\,dt = \int_s^\infty e^{-\lambda(t-s)}\mathbb{E}_{\widetilde{\mathbb{P}}^2}[h(y_t)]\,dt$. By the uniqueness of the Laplace transform together with the continuity of the mappings $t \mapsto \mathbb{E}_{\widetilde{\mathbb{P}}^i}[h(y_t)]$, $i = 1, 2$, we obtain $\mathbb{E}_{\widetilde{\mathbb{P}}^1}[h(y_t)] = \mathbb{E}_{\widetilde{\mathbb{P}}^2}[h(y_t)]$, $t \ge s$, if h is bounded continuous. By a monotone class argument, the previous equality also extends to bounded measurable functions. Hence, $\widetilde{\mathbb{P}}^1_{s,x}(t) = \widetilde{\mathbb{P}}^2_{s,x}(t)$ and pushing forward the previous equality with respect to the law μ of the initial condition, we get $\mathbb{P}^1(t) = \mathbb{P}^2(t)$ for all $t \ge s$.

Finally, since $\mathbb{P}^1$ and $\mathbb{P}^2$ share the same one-dimensional marginal distribution, we remark that $\mathbb{P}^1$ and $\mathbb{P}^2$ are two solutions of the same *standard* martingale problem associated to the following operator with time dependent coefficients: $\widehat{\mathcal{L}}_t := \widehat{b}_i(t,x)\partial_i + \frac{1}{2}\widehat{a}_{i,j}(t,x)\partial^2_{i,j}$ with $\widehat{b}(t,x) := b(t,x,\mathbb{P}_1(t)) = b(t,x,\mathbb{P}_2(t))$, $\widehat{\sigma}(t,x) := \sigma(t,x,\mathbb{P}_1(t)) = \sigma(t,x,\mathbb{P}_2(t))$, $\widehat{a} := \widehat{\sigma}\widehat{\sigma}^*$ for which the well-posedness follows from Stroock and Varadhan [SV79]. Hence, we conclude that the finite-dimensional distributions of $\mathbb{P}^1$ and $\mathbb{P}^2$ coincide so that $\mathbb{P}^1 = \mathbb{P}^2$ and the martingale problem associated to the SDE (1.1) is well-posed. The proof of Theorem 3.3 is now complete.

Remark 4.1. (Extension of Theorem 3.3 by an approximation argument) The previous result on weak existence and uniqueness for the SDE (1.1) can be slightly improved in the case of scalar interactions by means of an approximation argument that we now briefly explain without going into too much technical detail. Similar arguments can be used for other examples.

Assuming $d = q = N = 1$ for simplicity, that is, $b(t,x,\mu) = b(t,x,\int\psi(x')\mu(dx'))$ and $a(t,x,\mu) = \sigma^2(t,x,\int\varphi(x')\mu(dx'))$, it is possible to establish the well-posedness of the martingale problem under the following weaker assumption: $(t,x,z) \mapsto b(t,x,z)$ and σ are bounded and continuous functions, ψ is bounded measurable, $x \mapsto \sigma(t,x,z)$ and φ are η-Hölder, $z \mapsto \sigma(t,x,z)$ is continuously differentiable and σ^2 is uniformly elliptic.

From Theorem 174 p.111 of Kestelman [Kes60], there exists a sequence $(\psi_N)_{N\ge1}$ of continuous functions defined on $\mathbb{R}$ such that $\sup_{N\ge1}|\psi_N|_\infty \le |\psi|_\infty$ and $\lim_{N\uparrow\infty}\psi_N = \psi$ a.e. One may also approximate σ by a sequence $(\sigma_N)_{N\ge1}$ such that $\lim_N a_N(t,x,z) = a(t,x,z)$ for every t,x,z, $z \mapsto a_N(t,x,z)$ is $\mathcal{C}^\infty$ and satisfying $\sup_{N\ge1}\sup_{t,x,z}|\partial_z\sigma_N(t,x,z)| \le \sup_{t,x,z}|\partial_z\sigma(t,x,z)|_\infty$, $\sup_{N\ge1}[a_N]_H \le [a]_H$.

Weak existence and uniqueness to the SDE (1.1) with coefficients $b_N(t,x,\mu) = b(t,x,\int \psi_N(x')\mu(dx'))$ and $a_N(t,x,\mu) = \sigma_N^2(t,x,\int \varphi(x')\mu(dx'))$ defined above, follow from Theorem 3.3. Denote by $(\mathbb{P}^N)_{N\geq 1}$ the associated sequence of probability measures. From the boundedness of the coefficients, the sequence $(\mathbb{P}^N)_{N\geq 1}$ is tight and up to an extraction of a subsequence it converges to a probability measure $\mathbb{P}^\infty$. Passing to the limit in the characterisation of the martingale problem in terms of conditional expectations, it follows that $\mathbb{P}^\infty$ is a solution to the martingale.

Then, the key idea consists in passing to the limit as $N \uparrow \infty$ in the relation (4.53) in order to prove that it still holds for any weak solution of the limit equation. We proceed as follows and deliberately skip some technical details.

We first remark that the transition density $p_N(\mu,s,t,x,z)$ associated to the SDE with coefficients b_N and σ_N exists and can be represented in terms of an infinite series as in (3.10) but with b and a replaced by b_N and a_N in $\mathcal{H}$ and $\widehat{p}$. By dominated convergence, one may pass to the limit as $N \uparrow \infty$ in the representation so that $p_N(\mu,s,t,x,z) \to p_\infty(\mu,s,t,x,z)$ where $p_\infty(\mu,s,t,x,z)$ satisfies a similar representation in infinite series but with coefficients $(t,x) \mapsto b(t,x,\mathbb{P}^\infty(t))$ and $(t,x) \mapsto a(t,x,\mathbb{P}^\infty(t))$ in the very definition of $\widehat{p}$ and of the kernel $\mathcal{H}$. Hence, by (3.11) and dominated convergence,

$$\begin{aligned}\lim_N S_\lambda^N h(s,x,\mu) &:= \lim_N \int_s^\infty e^{-\lambda(t-s)} \int h(z) p_N(\mu,s,t,x,z)\,dz\,dt \\ &= \int_s^\infty e^{-\lambda(t-s)} \int h(z) p_\infty(\mu,s,t,x,z)\,dz\,dt \\ &=: S_\lambda^\infty h(s,x,\mu).\end{aligned}$$

Then, we importantly note that, from Proposition 4.1, $(s,x,\mu) \mapsto \bar{p}_N(\mu,s,t,x,z) \in \mathcal{C}^{1,2,2}([0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ and satisfies the estimates (4.7) to (4.9) with constants C, c that do not depend on N. Hence, relabelling the indices if necessary, noting that by weak uniqueness of the proxy process (4.1) with coefficients a (see Lemma 4.1), $\mu \mapsto \bar{p}_N(\mu,s,t,x,z)$ converges to $\mu \mapsto \bar{p}(\mu,s,t,x,z)$, from the Arzelà-Ascoli theorem, we may assert that $\{v \mapsto \partial_v^n[\partial_\mu \bar{p}_N(\mu,s,t,x,z)](v), N \geq 1\}$ converges uniformly so that $v \mapsto \partial_v^n[\partial_\mu \bar{p}(\mu,s,t,x,z)(v)]$ is continuously differentiable for $n=0$ and continuous for $n=1$. Hence, passing to the limit as $N \uparrow \infty$, by dominated convergence, we get

$$\lim_N S_\lambda^N \left(\int_s^\infty e^{-\lambda(t-s)} \mathcal{S}_{s,t}^N h\,dt \right) = S_\lambda^\infty \left(\int_s^\infty e^{-\lambda(t-s)} \mathcal{S}_{s,t} h\,dt \right)$$

where $\mathcal{S}_{s,t}^N$ is defined as in (4.48) and (4.49) but with coefficients b_N and a_N instead of b and a. Finally, by dominated convergence,

$$\begin{aligned}\lim_N \widehat{R}_\lambda^N h(s,x,\mu) &:= \lim_N \int_s^\infty e^{-\lambda(t-s)} \int h(y) \bar{p}_N^y(\mu,s,t,x,y)\,dy \\ &= \int_s^\infty e^{-\lambda(t-s)} \int h(y) \bar{p}^y(\mu,s,t,x,y)\,dy \\ &=: \widehat{R}_\lambda h(s,x,\mu).\end{aligned}$$

By passing to the limit in (4.53), we thus conclude that for any weak solution $\mathbb{P}^\infty$ of the SDE (1.1) with coefficients b and σ

$$S_\lambda^\infty h - S_\lambda^\infty \left(\int_s^\infty e^{-\lambda(t-s)} \mathcal{S}_{s,t} h\,dt \right) = \widehat{R}_\lambda h$$

and one completes the proof of weak uniqueness by following the same lines of reasonings as those employed in the rest of the proof of Theorem 3.3.

5. Existence and regularity properties of the transition density

This section is dedicated to the proof of Theorem 3.9. Hence, throughout this section, we assume that (**HC**), (**HR**$_+$), (**HE**) are in force and that both maps $\mu \mapsto b(s,x,\mu)$, $a(s,x,\mu)$ belong to (**CS**$_+$).

5.1. **Strategy of proof.** Our strategy is essentially the same as the one employed for the proof of Proposition 4.1. To be more specific, for a given initial condition $(s,\mu) \in \mathbb{R}_+ \times \mathcal{P}_2(\mathbb{R}^d)$ and a probability measure $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, $\nu \neq \mu$, we let $\mathbb{P}^{(0)} = (\mathbb{P}^{(0)}(t))_{t\geq s}$ be the probability measure on $\mathcal{C}([s,\infty),\mathbb{R}^d)$, endowed with its canonical filtration, satisfying $\mathbb{P}^{(0)}(t) = \nu$, $t \geq s$, and we consider the following recursive sequence of probability measures $\{\mathbb{P}^{(m)}; m \geq 0\}$, with time marginals $(\mathbb{P}^{(m)}(t))_{t\geq s}$, where, $\mathbb{P}^{(m)}$ being given, $\mathbb{P}^{(m+1)}$ is the unique solution to the following martingale problem

(i) $\mathbb{P}^{(m+1)}(y(r) \in \Gamma; 0 \leq r \leq s) = \mu(\Gamma)$, for all $\Gamma \in \mathcal{B}(\mathbb{R}^d)$.

(ii) For all $f \in \mathcal{C}_b^2(\mathbb{R}^d)$,

$$f(y_t) - f(y_s) - \int_s^t \left\{ \sum_{i=1}^d b_i(r, y_r, \mathbb{P}^{(m)}(r))\partial_i f(y_r) + \sum_{i,j=1}^d \frac{1}{2} a_{i,j}(r, y_r, \mathbb{P}^{(m)}(r))\partial_{i,j}^2 f(y_r) \right\} dr$$

is a continuous square-integrable martingale under $\mathbb{P}^{(m+1)}$.

Note that, under the considered assumptions, the well-posedness of the above martingale problem follows from standard results, see e.g. [SV79], so that there exists a unique weak solution to the SDE

$$X_t^{s,\xi,(m+1)} = \xi + \int_s^t b(r, X_r^{s,\xi,(m+1)}, [X_r^{s,\xi,(m)}])dr + \int_s^t \sigma(r, X_r^{s,\xi,(m+1)}, [X_r^{s,\xi,(m)}])dW_r. \tag{5.1}$$

We also associate to the above dynamics the decoupled stochastic flow given by the unique weak solution to SDE

$$X_t^{s,x,\mu,(m+1)} = x + \int_s^t b(r, X_r^{s,x,\mu,(m+1)}, [X_r^{s,\xi,(m)}])dr + \int_s^t \sigma(r, X_r^{s,x,\mu,(m+1)}, [X_r^{s,\xi,(m)}])dW_r. \tag{5.2}$$

We point out that the notation $X_t^{s,x,\mu,(m+1)}$ makes sense since by weak uniqueness of solution to the SDE (5.1), the law $[X_t^{s,\xi,(m)}]$ only depends on the initial condition ξ through its law μ.

From [Fri64], for any $m \geq 0$, the two random variables $X_t^{s,\xi,(m)}$ and $X_t^{s,x,\mu,(m)}$ admit a density respectively denoted by $p_m(\mu, s, t, z)$ and $p_m(\mu, s, t, x, z)$. Moreover, the following relation is satisfied

$$\forall z \in \mathbb{R}^d, \quad p_m(\mu, s, t, z) = \int p_m(\mu, s, t, x, z)\mu(dx) \tag{5.3}$$

where for all $m \geq 1$

$$p_m(\mu, s, t, x, z) = \sum_{k \geq 0} (\widehat{p}_m \otimes \mathcal{H}_m^{(k)})(\mu, s, t, x, z), \tag{5.4}$$

with

$$\begin{aligned}
\widehat{p}_m(\mu, s, r, t, x, z) &= g\left(\int_r^t a(v, z, [X_v^{s,\xi,(m-1)}])dv, z - x \right), \\
\mathcal{H}_m(\mu, s, r, t, x, z) &= \left\{ -\sum_{i=1}^d b_i(r, x, [X_r^{s,\xi,(m-1)}]) H_1^i \left(\int_r^t a(v, z, [X_v^{s,\xi,(m-1)}])dv, z - x \right) \right. \\
&\quad + \frac{1}{2}(a_{i,j}(r, x, [X_r^{s,\xi,(m-1)}]) - a_{i,j}(r, z, [X_r^{s,\xi,(m-1)}])) \\
&\quad \left. \times H_2^{i,j} \left(\int_r^t a(v, z, [X_v^{s,\xi,(m-1)}])dv, z - x \right) \right\} \widehat{p}_m(\mu, s, r, t, x, z)
\end{aligned}$$

and $\mathcal{H}_m^{(k+1)}(\mu, s, t, x, z) = (\mathcal{H}_m^{(k)} \otimes \mathcal{H}_m)(\mu, s, t, x, z)$, $\mathcal{H}_m^{(0)} = I_d$, with the convention that $[X_t^{s,\xi,(0)}] = \mathbb{P}^{(0)}(t) = \nu$, $t \geq 0$. In what follows, we will often make use of the following estimates: there exist constant $c := c(\lambda) > 1$, $C := C(T, a, b, \eta) > 0$, such that for all $0 \leq s < t \leq T$, for all integer k, one has

$$\forall (x, z) \in (\mathbb{R}^d)^2, \quad |\widehat{p}_m \otimes \mathcal{H}_m^{(k)}(\mu, s, t, x, z)| \leq C^k (t-s)^{k\frac{\eta}{2}} \prod_{i=1}^k B\left(1 + \frac{(i-1)\eta}{2}, \frac{\eta}{2}\right) g(c(t-s), z - x) \tag{5.5}$$

where $B(k, \ell) = \int_0^1 (1-v)^{-1+k} v^{-1+\ell} dv$ stands for the Beta function. As a consequence, from the asymptotics of the Beta function, the series (5.4) converge absolutely and uniformly for $(\mu, x, z) \in \mathcal{P}_2(\mathbb{R}^d) \times (\mathbb{R}^d)^2$ and satisfies: for all $m \geq 1$, for any $0 \leq s < t \leq T$ and any $(\mu, x, z) \in \mathcal{P}_2(\mathbb{R}^d) \times (\mathbb{R}^d)^2$

$$|\partial_x^n p_m(\mu, s, t, x, z)| \leq C(t-s)^{-\frac{n}{2}} g(c(t-s), z - x), \; n = 0, 1, 2, \tag{5.6}$$

where $C := C(T, a, b, \eta)$ and $c := c(\lambda)$ are two positive constants and for all $(x, x') \in (\mathbb{R}^d)^2$, for any $m \geq 1$,

$$\begin{aligned}
&|\partial_x^n p_m(\mu, s, r, x, z) - \partial_x^n p_m(\mu, s, r, x', z)| \\
&\qquad \leq C \frac{|x - x'|^\beta}{(r-s)^{\frac{n+\beta}{2}}} \left\{ g(c(r-s), z - x) + g(c(r-s), z - x') \right\}, \; n = 0, 1, 2,
\end{aligned} \tag{5.7}$$

where $\beta \in [0, 1]$ if $n = 0$, 1 and $\beta \in [0, \eta)$ if $n = 2$. We refer to Friedman [Fri64] for a proof of the above estimate.

Denote by $\Phi_m(\mu,s,r,t,x_1,x_2)$ the solution to the Volterra integral equation

$$\Phi_m(\mu,s,r,t,x_1,x_2)=\mathcal{H}_m(\mu,s,r,t,x_1,x_2)+(\mathcal{H}_m\otimes\Phi_m)(\mu,s,r,t,x_1,x_2). \tag{5.8}$$

From the space-time inequality (1.4), it is easily seen that the singular kernel $\mathcal{H}_m(\mu,s,r,t,x_1,x_2)$ induces an integrable singularity in time in the above space-time convolution so that the solution exists and is given by the (uniform) convergent series

$$\Phi_m(\mu,s,r,t,x_1,x_2)=\sum_{k\geq 1}\mathcal{H}_m^{(k)}(\mu,s,r,t,x_1,x_2) \tag{5.9}$$

so that (5.4) now writes

$$p_m(\mu,s,t,x,z)=\widehat{p}_m(\mu,s,t,x,z)+\int_s^t\int_{\mathbb{R}^d}\widehat{p}_m(\mu,s,r,x,y)\Phi_m(\mu,s,r,t,y,z)\,dy\,dr. \tag{5.10}$$

Moreover, from Theorem 7, Chapter 1 in [Fri64], for any $m\geq 1$, the map $x\mapsto\Phi_m(\mu,s,r,t,x,z)$ is Hölder-continuous. More precisely, for any $\beta\in[0,\eta)$, there exist two positive constants $C:=C(T,a,b,\eta,\lambda)$, $c(\lambda)$, thus do not depending on m, such that

$$\begin{aligned}&|\Phi_m(\mu,s,r,t,x,z)-\Phi_m(\mu,s,r,t,y,z)|\\&\qquad\leq C\frac{|x-y|^\beta}{(t-r)^{1+\frac{\beta-\eta}{2}}}\left\{g(c(t-r),z-x)+g(c(t-r),z-y)\right\}.\end{aligned} \tag{5.11}$$

With the above notations and properties, we prove the following key proposition whose proof is postponed to the next subsection.

Proposition 5.1. Let $T>0$. For any fixed $(t,z)\in(0,T]\times\mathbb{R}^d$, for all $m\geq 1$, the following properties hold:

- The mapping $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto p_m(\mu,s,t,x,z)$ is in $\mathcal{C}^{1,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.
- There exist two positive constants $C:=C(|b|_\infty,[b]_H,|\widetilde{b}|_\infty,[\widetilde{b}]_H,|a|_\infty,,[a]_H,|\widetilde{a}|_\infty,[\widetilde{a}]_H,T)$, $c:=c(\lambda)$, thus do not depending on m, such that for any $(\mu,s,x,x',z,v,v')\in\mathcal{P}_2(\mathbb{R}^d)\times[0,t)\times(\mathbb{R}^d)^5$,

$$|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](v)|\leq\frac{C}{(t-s)^{\frac{1+n-\eta}{2}}}\,g(c(t-s),z-x),\quad n=0,1, \tag{5.12}$$

$$|\partial_s p_m(\mu,s,t,x,z)|\leq\frac{C}{t-s}\,g(c(t-s),z-x), \tag{5.13}$$

$$\begin{aligned}&|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](v)-\partial_v^n[\partial_\mu p_m(\mu,s,t,x',z)](v)|\\&\qquad\leq C\frac{|x-x'|^\beta}{(t-s)^{\frac{1+n+\beta-\eta}{2}}}\left\{g(c(t-s),z-x)+g(c(t-s),z-x')\right\},\end{aligned} \tag{5.14}$$

where $\beta\in[0,1]$ for $n=0$ and $\beta\in[0,\eta)$ for $n=1$,

$$\begin{aligned}&\forall\beta\in[0,\eta),\quad|\partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v)-\partial_v[\partial_\mu p_m(\mu,s,t,x,z)](v')|\\&\qquad\leq C\frac{|v-v'|^\beta}{(t-s)^{1+\frac{\beta-\eta}{2}}}\,g(c(t-s),z-x).\end{aligned} \tag{5.15}$$

There exist $C:=C(|b|_\infty,[b]_H,|\widetilde{b}|_\infty,[\widetilde{b}]_H,|\widetilde{B}|_\infty,[\widetilde{B}]_H,|a|_\infty,[a]_H,|\widetilde{a}|_\infty,[\widetilde{a}]_H,|\widetilde{A}|_\infty,[\widetilde{A}]_H,T)>0$, $c:=c(\lambda)>0$, thus do not depending on m, such that for any $(\mu,\mu',s,x,z,v)\in(\mathcal{P}_2(\mathbb{R}^d))^2\times[0,t)\times(\mathbb{R}^d)^3$,

$$|\partial_x^n p_m(\mu,s,t,x,z)-\partial_x^n p_m(\mu',s,t,x,z)]|\leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{n+\beta-\eta}{2}}}\,g(c(t-s),z-x), \tag{5.16}$$

where $\beta\in[0,1]$ for $n=0$, 1 and $\beta\in[0,\eta)$ for $n=2$,

$$\begin{aligned}&|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](v)-\partial_v^n[\partial_\mu p_m(\mu',s,t,x,z)](v)|\\&\qquad\leq C\frac{W_2^\beta(\mu,\mu')}{(t-s)^{\frac{1+n+\beta-\eta}{2}}}\,g(c(t-s),z-x),\end{aligned} \tag{5.17}$$

where $\beta\in[0,1]$ for $n=0$ and $\beta\in[0,\eta)$ for $n=1$, and for all $(s_1,s_2)\in[0,t)^2$,

$$
\begin{aligned}
&|\partial_x^n p_m(\mu,s_1,t,x,z)-\partial_x^n p_m(\mu,s_2,t,x,z)|\\
&\quad\le C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{n}{2}+\beta}}\,g(c(t-s_1),z-x)+\frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{n}{2}+\beta}}\,g(c(t-s_2),z-x)\right\},
\end{aligned}\tag{5.18}
$$

where $\beta\in[0,1]$ for $n=0$, $\beta\in[0,\frac{1+\eta}{2})$ for $n=1$ and $\beta\in[0,\frac{\eta}{2})$ for $n=2$ and

$$
\begin{aligned}
&|\partial_v^n[\partial_\mu p_m(\mu,s_1,t,x,z)](v)-\partial_v^n[\partial_\mu p_m(\mu,s_2,t,x,z)](v)|\\
&\quad\le C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}}\,g(c(t-s_1),z-x)+\frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}}\,g(c(t-s_2),z-x)\right\},
\end{aligned}\tag{5.19}
$$

where $\beta\in[0,\frac{\eta}{2})$.

The proof of the above result being rather long and technical it is postponed to the subsection 5.2. Then, since the coefficients b_i, $a_{i,j}$ are bounded and the initial condition $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, the sequence $(\mathbb{P}^{(m)})_{m\ge0}$ is tight. Relabelling the indices if necessary, we may assert that $(\mathbb{P}^{(m)})_{m\ge0}$ converges weakly to a probability measure $\mathbb{P}^\infty$. From standard arguments that we omit (passing to the limit in the characterisation of the martingale problem solved by $\mathbb{P}^{(m)}$) we deduce that $\mathbb{P}^\infty$ is the probability measure $\mathbb{P}$ induced by the unique weak solution to the McKean-Vlasov SDE (1.1). As a consequence, every convergent subsequence converges to the same limit $\mathbb{P}$ and so does the original sequence $(\mathbb{P}^{(m)})_{m\ge1}$.

By Lebesgue's dominated convergence, using (5.5), one may pass to the limit as $m\uparrow\infty$ in the parametrix infinite series (5.4) and thus deduce that the sequence $\{p_m(\mu,s,t,x,z),\,m\ge1\}$ converges to $p(\mu,s,t,x,z)$ given by the infinite series (3.10) for any fixed $(\mu,s,t,x,z)\in\mathcal{P}_2(\mathbb{R}^d)\times[0,T]^2\times(\mathbb{R}^d)^2$ satisfying $0\le s<t\le T$.

Finally, the same lines of reasoning as those employed for the second and third steps in the proof of Proposition 4.1 apply. To be more specific, combining (5.7), the estimates (5.12) to (5.19) with Arzelà-Ascoli's theorem yield that the mapping $(s,x,\mu)\mapsto p(\mu,s,t,x,z)$ is $\mathcal{C}^{0,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$ and passing to the limit along the corresponding subsequence in the above estimates allows to conclude that the estimates (3.15) and (3.18) to (3.23) are valid. One then deduces that $(s,x,\mu)\mapsto p(\mu,s,t,x,z)\in\mathcal{C}^{1,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$ by combining the Markov property of the unique solution to the SDE (1.1), inherited from the well-posedness of the related martingale problem, with Proposition 2.1 applied to $p([X_s^{s-h,\xi}],s,t,X_s^{s-h,x,\mu},z)$ on the interval $[s-h,s]$. In particular, by a similar argument as the one employed in the third step of the proof of Proposition 4.1, one proves that the map $(s,x,\mu)\mapsto p(\mu,s,t,x,z)$ is a solution to the following PDE

$$
\partial_s p(\mu,s,t,x,z)=-\mathcal{L}_s p(\mu,s,t,x,z)\text{ on }[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)
$$

with the terminal condition: $\lim_{s\uparrow t}p(\mu,s,t,x,z)=\delta_z(x)$ in the weak sense. Finally, the estimates (3.16) and (3.17) follow from the previous identity combined with the estimates (5.6) and (3.15) in the one hand and the estimates (3.13) and (3.18) on the other hand. We now move to the proof of Proposition 5.1.

5.2. **Proof of Proposition 5.1.** We proceed by induction on m. For $m=1$, observe that $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto p_1(\mu,s,t,x,z)=\sum_{k\ge0}(\widehat{p}_1\otimes\mathcal{H}_1^{(k)})(\mu,s,t,x,z)$ is continuous, where we emphasise from the very definition of our iterative scheme that $\widehat{p}_1$ and H_1 do not depend on the law μ but only on the initial probability measure $\mathbb{P}^{(0)}$ of the iterative scheme. Hence, $\mu\mapsto p_1(\mu,s,t,x,z)$ is continuously L-differentiable with $\partial_\mu p_1(\mu,s,t,x,z)(v)=\partial_v[\partial_\mu p_1(\mu,s,t,x,z)](v)\equiv0$. For any $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, $(s,x)\mapsto p_1(\mu,s,t,x,z)\in\mathcal{C}^{1,2}([0,t)\times\mathbb{R}^d)$.
Moreover, differentiating n-times ($n=1,2$) w.r.t the variable x the relation

$$
\begin{aligned}
\int_{\mathbb{R}^d}\widehat{p}_m(\mu,s,r,x,y)\Phi_m(\mu,s,r,t,y,z)\,dy&=\int_{\mathbb{R}^d}\widehat{p}_m(\mu,s,r,x,y)[\Phi_m(\mu,s,r,t,y,z)-\Phi_m(\mu,s,r,t,x',z)]\,dy\\
&\quad+\Phi_m(\mu,s,r,t,x',z)\int_{\mathbb{R}^d}[\widehat{p}_m(\mu,s,r,x,y)-\widehat{p}_m^{x'}(\mu,s,r,x,y)]\,dy\\
&\quad+\Phi_m(\mu,s,r,t,x',z)
\end{aligned}
$$

and then choosing $x'=x$, by Lebesgue's differentiation theorem, we get

$$
\begin{aligned}
\int_{\mathbb{R}^d}\partial_x^n\widehat{p}_m(\mu,s,r,x,y)\Phi_m(\mu,s,r,t,y,z)\,dy&=\int_{\mathbb{R}^d}\partial_x^n\widehat{p}_m(\mu,s,r,x,y)[\Phi_m(\mu,s,r,t,y,z)-\Phi_m(\mu,s,r,t,x,z)]\,dy\\
&\quad+\Phi_m(\mu,s,r,t,x,z)\int_{\mathbb{R}^d}[\partial_x^n\widehat{p}_m^y(\mu,s,r,x,y)-\partial_x^n\widehat{p}_m^x(\mu,s,r,x,y)]\,dy.
\end{aligned}
$$

The η-Hölder regularity of $x \mapsto a(t,x,\mu)$ and the space-time inequality (1.4) implies

(5.20)
$$|\partial_x^n \widehat{p}_m^y(\mu,s,r,x,y) - \partial_x^n \widehat{p}_m^x(\mu,s,r,x,y)| \leq K|y-x|^\eta (r-s)^{-\frac{n}{2}} g(c(r-s),y-x) \leq K(r-s)^{\frac{-n+\eta}{2}} g(c(r-s),y-x)$$

and, from (5.11) and the space-time inequality (1.4),

$$\Big|\int_{\mathbb{R}^d} \partial_x^n \widehat{p}_m(\mu,s,r,x,y)[\Phi_m(\mu,s,r,t,y,z) - \Phi_m(\mu,s,r,t,x,z)]\,dy\Big| \leq K(t-r)^{-1+\frac{\eta-\beta}{2}}(r-s)^{\frac{-n+\beta}{2}} g(c(t-r),z-x)$$

for any $\beta \in [0,\eta)$. Hence, differentiating (5.10) w.r.t the variable x, from Lebesgue's differentiation theorem, we obtain

$$\begin{aligned}
\partial_x^n p_m(\mu,s,t,x,z) &= \partial_x^n \widehat{p}_m(\mu,s,t,x,z) \\
&+ \int_s^t \int_{\mathbb{R}^d} \partial_x^n \widehat{p}_m(\mu,s,r,x,y)\Big[\Phi_m(\mu,s,r,t,y,z) - \Phi_m(\mu,s,r,t,x,z)\Big]\,dy\,dr \\
&+ \int_s^t \Phi_m(\mu,s,r,t,x,z) \int_{\mathbb{R}^d} \Big[\partial_x^n \widehat{p}_m^y(\mu,s,r,x,y) - \partial_x^n \widehat{p}_m^x(\mu,s,r,x,y)\Big]\,dy\,dr,\ n=1,2.
\end{aligned} \tag{5.21}$$

Observe that since $[X_v^{s,\xi,(0)}] = \nu$ we have that $\Phi_1(\mu,s,r,t,x,z) \equiv \Phi_1(r,t,x,z)$ and $\widehat{p}_1(\mu,s,s,t,x,z) \equiv \widehat{p}_1(s,t,x,z)$. In a completely analogous manner, one may differentiate w.r.t the variable s the relation (5.10) for $m=1$. We obtain

$$\begin{aligned}
\partial_s p_1(\mu,s,t,x,z) &= \partial_s \widehat{p}_1(s,t,x,z) - \Phi_1(s,t,x,z) \\
&+ \int_s^t \int_{\mathbb{R}^d} \partial_s \widehat{p}_1(s,r,x,y)\Big[\Phi_1(r,t,y,z) - \Phi_1(r,t,x,z)\Big]\,dy\,dr \\
&+ \int_s^t \Phi_1(r,t,x,z) \int_{\mathbb{R}^d} \Big[\partial_s \widehat{p}_1^y(s,r,x,y) - \partial_s \widehat{p}_1^x(s,r,x,y)\Big]\,dy\,dr.
\end{aligned} \tag{5.22}$$

Then, Lebesgue's dominated convergence theorem, the inequality $|\partial_s \widehat{p}_1^y(s,r,x,y) - \partial_s \widehat{p}_1^x(s,r,x,y)| \leq C|x-y|^\eta (r-s)^{-1} g(c(r-s),y-x) \leq C(r-s)^{-1+\frac{\eta}{2}} g(c(r-s),y-x)$, derived from the space-time inequality (1.4), (5.11) as well as the continuity of the mappings $(s,x,\mu) \mapsto \Phi_1(r,t,x,z)$, $a(s,x,\mu)$, $b(s,x,\mu)$, allow to conclude that the three maps $[0,t) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (s,x,\mu) \mapsto \partial_x p_1(\mu,s,t,x,z)$, $\partial_x^2 p_1(\mu,s,t,x,z)$ are continuous. From the previous computations, the estimates (5.12) to (5.17) and (5.19) are straightforward for $m=1$.

In order to derive the estimate (5.18), we proceed as follows. First, observe that we may assume without loss of generality that $|s_1 - s_2| \leq t - s_1 \vee s_2$. Indeed, if $|s_1 - s_2| \geq t - s_1 \vee s_2$, then from (5.6) one directly gets

$$\begin{aligned}
&|\partial_x^n p_1(\mu,s_1,t,x,z) - \partial_x^n p_1(\mu,s_2,t,x,z)| \\
&\quad = |\partial_x^n p_1(\mu,s_1 \vee s_2,t,x,z) - \partial_x^n p_1(\mu,s_2 \wedge s_2,t,x,z)| \\
&\quad \leq C\left\{\frac{1}{(t-s_1\vee s_2)^{\frac{n}{2}}} g(c(t-s_1\vee s_2),z-x) + \frac{1}{(t-s_1\wedge s_2)^{\frac{n}{2}}} g(c(t-s_1\wedge s_2),z-x)\right\} \\
&\quad \leq C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{n}{2}+\beta}} g(c(t-s_1\vee s_2),z-x) + \frac{(t-s_1\wedge s_2)^\beta}{(t-s_1\wedge s_2)^{\frac{n}{2}+\beta}} g(c(t-s_1\wedge s_2),z-x)\right\} \\
&\quad \leq C\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{n}{2}+\beta}} g(c(t-s_1),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{n}{2}+\beta}} g(c(t-s_2),z-x)\right\},
\end{aligned}$$

for any $\beta \in [0,1]$ which is the desired bound. Hence, for the rest of the proof of (5.18), we will assume that $|s_1 - s_2| \leq t - s_1 \vee s_2$.

From (5.21), for all $m \geq 1$, we easily obtain the following decomposition

$$\begin{aligned}
\partial_x^n p_m(\mu,s,t,x,z) &= \partial_x^n \widehat{p}_m(\mu,s,t,x,z) \\
&+ \int_s^{\frac{t+s}{2}} \int_{\mathbb{R}^d} \partial_x^n \widehat{p}_m(\mu,s,r,x,y)\Big[\Phi_m(\mu,s,r,t,y,z) - \Phi_m(\mu,s,r,t,x,z)\Big]\,dy\,dr \\
&+ \int_s^{\frac{t+s}{2}} \Phi_m(\mu,s,r,t,x,z) \int_{\mathbb{R}^d} \Big[\partial_x^n \widehat{p}_m^y(\mu,s,r,x,y) - \partial_x^n \widehat{p}_m^x(\mu,s,r,x,y)\Big]\,dy\,dr \\
&+ \int_{\frac{t+s}{2}}^t \int_{\mathbb{R}^d} \partial_x^n \widehat{p}_m(\mu,s,r,x,y)\Phi_m(\mu,r,t,y,z)\,dy\,dr
\end{aligned} \tag{5.23}$$

which in turn implies

$$
\begin{aligned}
&\partial_x^n p_1(\mu, s_1\vee s_2, t, x, z) - \partial_x^n p_1(\mu, s_1\wedge s_2, t, x, z)\\
&\quad = \partial_x^n \widehat{p}_1(s_1\vee s_2, t, x, z) - \partial_x^n \widehat{p}_1(s_1\wedge s_2, t, x, z)\\
&\qquad + \int_{s_1\vee s_2}^{\frac{t+s_1\vee s_2}{2}} \int_{\mathbb{R}^d} [\partial_x^n \widehat{p}_1(s_1\vee s_2, r, x, y) - \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)]\\
&\qquad\quad .\Big[\Phi_1(r, t, y, z) - \Phi_1(r, t, x, z)\Big]\, dy\, dr\\
&\qquad - \int_{s_1\wedge s_2}^{s_1\vee s_2} \int_{\mathbb{R}^d} \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)\Big[\Phi_1(r, t, y, z) - \Phi_1(r, t, x, z)\Big]\, dy\, dr\\
&\qquad + \int_{\frac{t+s_1\wedge s_2}{2}}^{\frac{t+s_1\vee s_2}{2}} \int_{\mathbb{R}^d} \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)\Big[\Phi_1(r, t, y, z) - \Phi_1(r, t, x, z)\Big]\, dy\, dr\\
&\qquad + \int_{s_1\vee s_2}^{\frac{t+s_1\vee s_2}{2}} \Phi_1(r, t, x, z) \int_{\mathbb{R}^d} \Big[\partial_x^n \widehat{p}_1^{\,y}(s_1\vee s_2, r, x, y) - \partial_x^n \widehat{p}_1^{\,y}(s_1\wedge s_2, r, x, y)\Big]\, dy\, dr\\
&\qquad - \int_{s_1\wedge s_2}^{s_1\vee s_2} \Phi_1(r, t, x, z) \int_{\mathbb{R}^d} \Big[\partial_x^n \widehat{p}_1^{\,y}(s_1\wedge s_2, r, x, y) - \partial_x^n \widehat{p}_1^{\,x}(s_1\wedge s_2, r, x, y)\Big]\, dy\, dr\\
&\qquad + \int_{\frac{t+s_1\wedge s_2}{2}}^{\frac{t+s_1\vee s_2}{2}} \Phi_1(r, t, x, z) \int_{\mathbb{R}^d} \Big[\partial_x^n \widehat{p}_1^{\,y}(s_1\wedge s_2, r, x, y) - \partial_x^n \widehat{p}_1^{\,x}(s_1\wedge s_2, r, x, y)\Big]\, dy\, dr\\
&\qquad + \int_{\frac{t+s_1\vee s_2}{2}}^{t} \int_{\mathbb{R}^d} \Big[\partial_x^n \widehat{p}_1(s_1\vee s_2, r, x, y) - \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)\Big]\Phi_1(r, t, y, z)\, dy\, dr\\
&\qquad + \int_{\frac{t+s_1\wedge s_2}{2}}^{\frac{t+s_1\vee s_2}{2}} \int_{\mathbb{R}^d} \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)\Phi_1(r, t, y, z)\, dy\, dr\\
&\quad =: \sum_{i=1}^{9} A_i^n.
\end{aligned}
$$

From the mean-value theorem and the inequality $|s_1 - s_2| \le t - s_1\vee s_2$, similarly to (4.39), we obtain

$$
\begin{aligned}
\forall \beta\in[0,1],\quad |A_1^n| &= |\partial_x^n \widehat{p}_1(s_1, t, x, z) - \partial_x^n \widehat{p}_1(s_2, t, x, z)|\\
&\le C\Big[\frac{|s_1 - s_2|^\beta}{(t-s_1)^{\frac{n}{2}+\beta}} g(c(t-s_1), z-x) + \frac{|s_1 - s_2|^\beta}{(t-s_2)^{\frac{n}{2}+\beta}} g(c(t-s_2), z-x)\Big].
\end{aligned}
\tag{5.24}
$$

Combining (5.11), (5.24) and the space-time inequality (1.4), for all $\alpha\in[0,\eta)$, we get

$$
\begin{aligned}
|A_2^n| &\le \int_{\mathbb{R}^d} |\partial_x^n \widehat{p}_1(s_1\vee s_2, r, x, y) - \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)||\Phi_1(r, t, y, z) - \Phi_1(r, t, x, z)|\, dy\\
&\le \frac{C}{(t-r)^{1+\frac{\alpha-\eta}{2}}}\Big[\frac{|s_1 - s_2|^\beta}{(r-s_1)^{\frac{n}{2}+\beta-\frac{\alpha}{2}}} g(c(t-s_1), z-x) + \frac{|s_1 - s_2|^\beta}{(r-s_2)^{\frac{n}{2}+\beta-\frac{\alpha}{2}}} g(c(t-s_2), z-x)\Big],
\end{aligned}
$$

so that,

$$
|A_2^n| \le C\Big[\frac{|s_1 - s_2|^\beta}{(t-s_1)^{\frac{n}{2}+\beta-\frac{\eta}{2}}} g(c(t-s_1), z-x) + \frac{|s_1 - s_2|^\beta}{(t-s_2)^{\frac{n}{2}+\beta-\frac{\eta}{2}}} g(c(t-s_2), z-x)\Big],
$$

where $\beta\in[0,1]$ if $n=0$, $\beta\in[0,(1+\eta)/2)$ if $n=1$ and $\beta\in[0,\eta/2)$ if $n=2$. For A_3^n, similar arguments yield

$$
\begin{aligned}
\Big|\int_{\mathbb{R}^d} \partial_x^n \widehat{p}_1(s_1\wedge s_2, r, x, y)\Big[\Phi_1(r, t, y, z) - \Phi_1(r, t, x, z)\Big]\, dy\Big| &\le C\frac{1}{(r - s_1\wedge s_2)^{\frac{n-\alpha}{2}}}\frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}}\\
&\quad\times g(c(t - s_1\wedge s_2), z-x)
\end{aligned}
$$

where we used the inequality $(t-r)^{-1} \leq (t-s_1\wedge s_2)^{-1}(1+(r-s_1\wedge s_2)/(t-r)) \leq 2(t-s_1\wedge s_2)^{-1}$, for $r < s_1 \vee s_2$ as well as $|s_1-s_2| \leq t - s_1 \vee s_2$. Hence, for all $\alpha \in [0,\eta)$

$$\begin{aligned}
|A_3^n| &\leq C\int_{s_1\wedge s_2}^{s_1\vee s_2} \frac{1}{(r-s_1\wedge s_2)^{\frac{n-\alpha}{2}}}\frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}}\,dr\, g(c(t-s_1\wedge s_2),z-x)\\
&\leq C\frac{|s_1-s_2|^{1+\frac{\alpha-n}{2}}}{(t-s_1\wedge s_2)^{1+\frac{\alpha-\eta}{2}}}\, g(c(t-s_1\wedge s_2),z-x)\\
&\leq C\frac{|s_1-s_2|^{\beta}}{(t-s_1\wedge s_2)^{\frac{n-\eta}{2}+\beta}}\, g(c(t-s_1\wedge s_2),z-x).
\end{aligned}$$

For A_4^n, we remark that since $|s_1-s_2| \leq t - s_1\vee s_2$, from (5.11) with $\beta = 0$, one gets

$$\begin{aligned}
|A_4^n| &\leq C\int_{\frac{t+s_1\wedge s_2}{2}}^{\frac{t+s_1\vee s_2}{2}} \frac{1}{(r-s_1\wedge s_2)^{\frac{n}{2}}}\frac{1}{(t-r)^{1-\frac{\eta}{2}}}\,dr\, g(c(t-s_1\wedge s_2),z-x)\\
&\leq \frac{C}{(t-s_1\wedge s_2)^{\frac{n}{2}}}\int_{\frac{t+s_1\wedge s_2}{2}}^{\frac{t+s_1\vee s_2}{2}} \frac{1}{(t-r)^{1-\frac{\eta}{2}}}\,dr\, g(c(t-s_1\wedge s_2),z-x)\\
&\leq C\frac{|s_1-s_2|^{\alpha}}{(t-s_1\wedge s_2)^{\frac{n-\eta}{2}+\alpha}} g(c(t-s_1\wedge s_2),z-x)
\end{aligned}$$

for any $\alpha \in [0,1]$.

We proceed similarly for A_5^n. If $|s_1-s_2| \leq r - s_1\vee s_2$, we use the mean-value theorem, the uniform η-Hölder regularity of $x \mapsto a(s,x,\mu)$ and the space-time inequality (1.4) to get

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d}\Big[\partial_x^n\widehat{p}_1^y(s_1\vee s_2,r,x,y)-\partial_x^n\widehat{p}_1^y(s_1\wedge s_2,r,x,y)\Big]\,dy\Big|\\
&\qquad= \Big|\int_{\mathbb{R}^d}\Big[\partial_x^n\widehat{p}_1^y(s_1\vee s_2,r,x,y)-\partial_x^n\widehat{p}_1^x(s_1\vee s_2,r,x,y)\\
&\qquad\qquad -(\partial_x^n\widehat{p}_1^y(s_1\wedge s_2,r,x,y)-\partial_x^n\widehat{p}_1^x(s_1\wedge s_2,r,x,y))\Big]\,dy\Big|\\
&\qquad\leq C\int_0^1 \frac{|s_1-s_2|}{(r-(\lambda s_1+(1-\lambda)s_2)^{1+\frac{n-\eta}{2}}}\,d\lambda\\
&\qquad\leq C\frac{|s_1-s_2|^{\alpha}}{(r-s_1\vee s_2)^{\frac{n-\eta}{2}+\alpha}}
\end{aligned}$$

for any $\alpha \in [0,1]$. Otherwise, if $|s_1-s_2| \geq r - s_1\vee s_2$ then one directly gets

$$\begin{aligned}
&\Big|\int_{\mathbb{R}^d}\Big[\partial_x^n\widehat{p}_1^y(s_1\vee s_2,r,x,y)-\partial_x^n\widehat{p}_1^y(s_1\wedge s_2,r,x,y)\Big]\,dy\Big|\\
&\qquad= \Big|\int_{\mathbb{R}^d}\Big[\partial_x^n\widehat{p}_1^y(s_1\vee s_2,r,x,y)-\partial_x^n\widehat{p}_1^x(s_1\vee s_2,r,x,y)\\
&\qquad\qquad -(\partial_x^n\widehat{p}_1^y(s_1\wedge s_2,r,x,y)-\partial_x^n\widehat{p}_1^x(s_1\wedge s_2,r,x,y))\Big]\,dy\Big|\\
&\qquad\leq C\frac{|s_1-s_2|^{\alpha}}{(r-s_1\vee s_2)^{\frac{n-\eta}{2}+\alpha}}
\end{aligned}$$

for any $\alpha \in [0,1]$. Hence, we conclude

$$|A_5^n| \leq C\frac{|s_1-s_2|^{\alpha}}{(t-s_1\vee s_2)^{\frac{n}{2}+\alpha-\eta}}\, g(c(t-s_1\vee s_2),z-x),$$

with $\alpha \in [0,1]$ for $n=0$, $\alpha \in [0,(1+\eta)/2)$ for $n=1$ and $\alpha \in [0,\eta/2)$ for $n=2$. Similar arguments yield

$$|A_6^n| \leq C\frac{|s_1-s_2|^{1+\frac{\eta-n}{2}}}{(t-s_1\wedge s_2)^{1-\frac{\eta}{2}}}\, g(c(t-s_1\wedge s_2),z-x) \leq C\frac{|s_1-s_2|^{\alpha}}{(t-s_1\wedge s_2)^{\frac{n}{2}+\alpha-\eta}}\, g(c(t-s_1\wedge s_2),z-x)$$

and

$$|A_7^n| \leq C\frac{|s_1-s_2|^{\alpha}}{(t-s_1\wedge s_2)^{\frac{n}{2}+\alpha-\eta}}\, g(c(t-s_1\wedge s_2),z-x)$$

with $\alpha \in [0,1]$ for $n=0$, $\alpha \in [0, \frac{1+\eta}{2})$ for $n=1$ and $\alpha \in [0, \eta/2)$ for $n=2$. From (5.24), for any $\alpha \in [0,1]$ we get

$$\begin{aligned}|A_8^n| &\leq C \int_{\frac{t+s_1\vee s_2}{2}}^t \frac{1}{(t-r)^{1-\frac{\eta}{2}}} \Big[\frac{|s_1-s_2|^\alpha}{(r-s_1\vee s_2)^{\frac{n}{2}+\alpha}} g(c(t-s_1\vee s_2), z-x) \\ &\quad + \frac{|s_1-s_2|^\alpha}{(r-s_1\wedge s_2)^{\frac{n}{2}+\alpha}} g(c(t-s_1\wedge s_2), z-x) \Big] dr \\ &\leq C \Big[\frac{|s_1-s_2|^\alpha}{(t-s_1\vee s_2)^{\frac{n}{2}+\alpha-\frac{\eta}{2}}} g(c(t-s_1\vee s_2), z-x) \\ &\quad + \frac{|s_1-s_2|^\alpha}{(t-s_1\wedge s_2)^{\frac{n}{2}+\alpha-\frac{\eta}{2}}} g(c(t-s_1\wedge s_2), z-x) \Big].\end{aligned}$$

Finally, using computations similar as those employed before, for all $\alpha \in [0,1]$, we get

$$|A_9^n| \leq C \frac{|s_1-s_2|}{(t-s_1\wedge s_2)^{\frac{n}{2}}(t-s_1\vee s_2)^{1-\frac{\eta}{2}}} g(c(t-s_1\wedge s_2), z-x) \leq C \frac{|s_1-s_2|^\alpha}{(t-s_1\vee s_2)^{\frac{n}{2}+\alpha-\frac{\eta}{2}}} g(c(t-s_1\wedge s_2), z-x).$$

This last bound concludes the proof of (5.18) at step $m=1$.

Assuming that the induction hypothesis is valid at step m, we then remark that if $(s_n, x_n, \mu_n)_{n\geq 1}$ is a sequence of $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$ satisfying $\lim_n |s_n - s| = \lim_n |x_n - x| = \lim_n W_2(\mu_n,\mu) = 0$ for some $(s,x,\mu) \in [0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)$, then, from similar arguments as those used in the proof of Proposition 4.1, namely the decomposition (4.20), in a completely analogous manner we obtain $\lim_n W_2([X_t^{s_n,\xi_n,(m)}], [X_t^{s,\xi,(m)}]) = 0$, where $[\xi_n] = \mu_n$ and $[\xi] = \mu$, so that $\lim_n a(t,x_n,[X_t^{s_n,\xi_n,(m)}]) = a(t,x,[X_t^{s,\xi,(m)}])$ and $\lim_n b(t,x_n,[X_t^{s_n,\xi_n,(m)}]) = b(t,x,[X_t^{s,\xi,(m)}])$. From the representation in infinite series (5.4) and the Lebesgue dominated convergence theorem, we deduce that the map $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d) \ni (s,x,\mu) \mapsto p_{m+1}(\mu,s,t,x,z)$ is continuous.

We next apply Lemma 2.1 to the maps $(s,\mu) \mapsto a(r,x,[X_r^{s,\xi,(m)}])$, $b(r,x,[X_r^{s,\xi,(m)}])$. Note that from the estimate (5.12), (5.13) and (5.6), the map $[0,r)\times\mathcal{P}_2(\mathbb{R}^d)\times(s,x,\mu) \mapsto p_m(\mu,s,r,x,z)$ satisfies the conditions of Definition 2.3, in particular condition (2.6). We thus deduce that $(s,\mu) \mapsto a(r,x,[X_r^{s,\xi,(m)}])$, $b(r,x,[X_r^{s,\xi,(m)}]) \in \mathcal{C}^{1,2}([0,r)\times\mathcal{P}_2(\mathbb{R}^d))$ with derivatives satisfying

(5.25)

$$\partial_s a(r,x,[X_r^{s,\xi,(m)}]) = \int\int (\widetilde{a}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}(r,x,x',[X_r^{s,\xi,(m)}]))\, \partial_s p_m(\mu,s,r,x',y')\, dy'\, \mu(dx'),$$

$$\partial_v^n[\partial_\mu a(t,x,[X_t^{s,\xi,(m)}])](v) = \int (\widetilde{a}(r,x,y',[X_t^{s,\xi,(m)}]) - \widetilde{a}(t,x,v,[X_t^{s,\xi,(m)}]))\, \partial_x^{1+n} p_m(\mu,s,t,v,y')\, dy'$$

(5.26)

$$+ \int\int (\widetilde{a}(t,x,y',[X_t^{s,\xi,(m)}]) - \widetilde{a}(t,x,x',[X_t^{s,\xi,(m)}]))\, \partial_v^n[\partial_\mu p_m(\mu,s,t,x',y')](v)\, dy'\, \mu(dx')$$

and similarly

$$\partial_s b(r,x,[X_r^{s,\xi,(m)}]) = \int\int (\widetilde{b}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{b}(r,x,x',[X_r^{s,\xi,(m)}]))\, \partial_s p_m(\mu,s,r,x',y')\, dy'\, \mu(dx'),$$

(5.27)

$$\begin{aligned}\partial_v^n[\partial_\mu b(t,x,[X_t^{s,\xi,(m)}])](v) &= \int (\widetilde{b}(r,x,y',[X_t^{s,\xi,(m)}]) - \widetilde{b}(t,x,v,[X_t^{s,\xi,(m)}]))\, \partial_x^{1+n} p_m(\mu,s,t,v,y')\, dy' \\ &\quad + \int\int (\widetilde{b}(t,x,y',[X_t^{s,\xi,(m)}]) - \widetilde{b}(t,x,x',[X_t^{s,\xi,(m)}]))\, \partial_v^n[\partial_\mu p_m(\mu,s,t,x',y')](v)\, dy'\, \mu(dx').\end{aligned}$$

As a consequence, the maps $[0,r)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\mathbb{R}^d \ni (s,x,\mu,v) \mapsto \partial_s a(r,x,[X_r^{s,\xi,(m)}])$, $\partial_v^n[\partial_\mu a(r,x,[X_r^{s,\xi,(m)}])](v)$ and $[0,r]\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d \ni (s,\mu,v) \mapsto \partial_s b(r,x,[X_r^{s,\xi,(m)}])$, $\partial_v^n[\partial_\mu b(r,x,[X_r^{s,\xi,(m)}])](v)$ are continuous for $n=0,1$. Moreover, from (5.13) at step m and the space-time inequality (1.4), one also derives the following bound:

$$\begin{aligned}|\partial_s b(r,x,[X_r^{s,\xi,(m)}])| + |\partial_s a(r,x,[X_r^{s,\xi,(m)}])| &\leq K \int\int (|y'-x'|^\eta \wedge 1)|\partial_s p_m(\mu,s,r,x',y')|\, dy'\, \mu(dx') \\ &\leq K(r-s)^{-1+\frac{\eta}{2}}.\end{aligned} \tag{5.28}$$

Similarly, from the estimate (5.12) and the space-time inequality (1.4), we also obtain

$$
\begin{aligned}
&|\partial_v^n[\partial_\mu b(t,x,[X_t^{s,\xi,(m)}])](v)| + |\partial_v^n[\partial_\mu a(t,x,[X_t^{s,\xi,(m)}])](v)| \\
&\qquad \le K\left\{(t-s)^{-\frac{1+n-\eta}{2}} + \int\int (|y-x'|^\eta \wedge 1)|\partial_v^n[\partial_\mu p_m(\mu,s,t,x',y)](v)|\,\mu(dx')\,dy\right\} \qquad (5.29)\\
&\qquad \le K(t-s)^{-\frac{1+n-\eta}{2}} \qquad (5.30)
\end{aligned}
$$

for $n = 0, 1$. As a consequence, the mapping $[0,r) \ni s \mapsto \mathcal{H}_{m+1}(\mu,s,r,t,x,z)$ is continuously differentiable for any $(\mu,x,z) \in \mathcal{P}_2(\mathbb{R}^d)\times(\mathbb{R}^d)^2$ with:

$$
\begin{aligned}
\partial_s\mathcal{H}_{m+1}(\mu,s,r,t,x,z) &:= \Big[-\sum_{i=1}^d \partial_s b_i\left(r,x,[X_r^{s,\xi,(m)}]\right) H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)\\
&\quad - \sum_{i=1}^d b_i\left(r,x,[X_r^{s,\xi,(m)}]\right)\partial_s H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)\\
&\quad + \frac{1}{2}\sum_{i,j=1}^d (\partial_s a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - \partial_s a_{i,j}(r,z,[X_r^{s,\xi,(m)}]))\\
&\qquad \times H_2^{i,j}(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x)\\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \Big(a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}])\Big)\\
&\qquad \times \partial_s H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)\Big]\widehat{p}_{m+1}(\mu,s,r,t,x,z)\\
&\quad + \Big[-\sum_{i=1}^d b_i(r,x,[X_r^{s,\xi,(m)}])H_1^i(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x)\\
&\qquad + \frac{1}{2}\sum_{i,j=1}^d \Big(a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,x,[X_r^{s,\xi,(m)}])\Big)\\
&\qquad\quad \times H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)})\,dv, z-x\right)\Big]\partial_s\widehat{p}_{m+1}(\mu,s,r,t,x,z).
\end{aligned} \qquad (5.31)
$$

The previous expression with the previous continuity results also yield the continuity of the mapping $[0,r)\times\mathcal{P}_2(\mathbb{R}^d) \ni (s,\mu) \mapsto \partial_s\mathcal{H}_{m+1}(\mu,s,r,t,x,z)$. Moreover, from (5.25), using either the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(r,x,y,\mu)$ or the η-Hölder regularity of $y \mapsto \widetilde{a}_{i,j}(r,x,y,\mu)$ with (5.13), we get

$$
|\partial_s a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - \partial_s a_{i,j}(r,z,[X_r^{s,\xi,(m)}])| \le K\left\{\frac{|z-x|^\eta}{(r-s)} \wedge \frac{1}{(r-s)^{\frac{\eta}{2}}}\right\}
$$

so that, from (5.28) and the space-time inequality (1.4), we get the following bound

$$
|\partial_s\mathcal{H}_{m+1}(\mu,s,r,t,x,z)| \le K\Big[\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)} \wedge \frac{1}{(t-r)(r-s)^{1-\frac{\eta}{2}}}\Big]g(c(t-r),z-x), \qquad (5.32)
$$

for two positive constants $K := K(T,a,b,\lambda,\eta)$, $c := c(\lambda)$, independent of m. Then, standard computations based on the previous estimate imply the convergence of the series $\sum_{k\ge0}(\mathcal{H}_{m+1}^{(k)}\otimes\partial_s\mathcal{H}_{m+1})(\mu,s,r,t,x,z)$. Moreover, if we formally differentiate w.r.t the variable s the relation $\Phi_{m+1}(\mu,s,r,t,x,z) = \mathcal{H}_{m+1}(\mu,s,r,t,x,z) + (\mathcal{H}_{m+1}\otimes\Phi_{m+1})(\mu,s,r,t,x,z)$, we get

$$
\begin{aligned}
\partial_s\Phi_{m+1}(\mu,s,r,t,x,z) &= \partial_s\mathcal{H}_{m+1}(\mu,s,r,t,x,z) + (\partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1})(\mu,s,r,t,x,z)\\
&\quad + (\mathcal{H}_{m+1}\otimes\partial_s\Phi_{m+1})(\mu,s,r,t,x,z)
\end{aligned}
$$

so that, iterating the previous relation, we deduce that the map $[0,r) \ni s \mapsto \Phi_{m+1}(\mu,s,r,t,x,z)$ is continuously differentiable with

$$
\begin{aligned}
\partial_s\Phi_{m+1}(\mu,s,r,t,x,z) &= \sum_{k\ge0}\Big(\mathcal{H}_{m+1}^{(k)}\otimes[\partial_s\mathcal{H}_{m+1} + \partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}]\Big)(\mu,s,r,t,x,z)\\
&= (\partial_s\mathcal{H}_{m+1} + \partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1})(\mu,s,r,t,x,z)\\
&\quad + (\Phi_{m+1}\otimes[\partial_s\mathcal{H}_{m+1} + \partial_s\mathcal{H}_{m+1}\otimes\Phi_{m+1}])(\mu,s,r,t,x,z).
\end{aligned} \qquad (5.33)
$$

We importantly note that from the estimate (5.32) and the continuity of $(s,\mu)\mapsto \partial_s\mathcal{H}_{m+1}(\mu,s,r,t,x,z)$ each term appearing in the above expression makes sense and is continuous w.r.t. (s,μ). We thus deduce that $[0,t)\ni s\mapsto \Phi_{m+1}(\mu,s,t,x,z)$ is continuously differentiable and $\partial_s\Phi_{m+1}$ is globally continuous w.r.t (s,μ). The same conclusion holds true for $[0,t)\ni s\mapsto p_{m+1}(\mu,s,t,x,z)$. More precisely, differentiating the relation (5.10) with respect to the variable s, then plugging (5.33) and using relation (5.10) again, we succesively get

$$\begin{aligned}\partial_s p_{m+1}(\mu,s,t,x,z)&=\partial_s\widehat{p}_{m+1}(\mu,s,t,x,z)-\Phi_{m+1}(\mu,s,t,x,z)\\&\quad+\int_s^t\int_{\mathbb{R}^d}\partial_s\widehat{p}_{m+1}(\mu,s,r,x,y)\Phi_{m+1}(\mu,s,r,t,y,z)\,dy\,dr\\&\quad+\int_s^t\int_{\mathbb{R}^d}\widehat{p}_{m+1}(\mu,s,r,x,y)\partial_s\Phi_{m+1}(\mu,s,r,t,y,z)\,dy\,dr\\&=\partial_s\widehat{p}_{m+1}(\mu,s,t,x,z)-\Phi_{m+1}(\mu,s,t,x,z)+\partial_s\widehat{p}_{m+1}\otimes\Phi_{m+1}(\mu,s,t,x,z)\\&\quad+p_{m+1}\otimes\partial_s\mathcal{H}_{m+1}(\mu,s,t,x,z)+[(p_{m+1}\otimes\partial_s\mathcal{H}_{m+1})\otimes\Phi_{m+1}](\mu,s,t,x,z).\end{aligned}\tag{5.34}$$

Again, we note that each term appearing in the above equality makes sense and is continuous w.r.t (s,x,μ). Indeed, using a similar decomposition as the one employed in (5.22), one gets

$$\begin{aligned}\partial_s\widehat{p}_{m+1}\otimes\Phi_{m+1}(\mu,s,t,x,z)&=\int_s^t\int_{\mathbb{R}^d}\partial_s\widehat{p}_{m+1}(\mu,s,r,x,y)\Big[\Phi_{m+1}(\mu,s,r,t,y,z)-\Phi_{m+1}(\mu,s,r,t,x,z)\Big]\,dy\,dr\\&\quad+\int_s^t\Phi_{m+1}(\mu,s,r,t,x,z)\int_{\mathbb{R}^d}\Big[\partial_s\widehat{p}^{\,y}_{m+1}(\mu,s,r,x,y)-\partial_s\widehat{p}^{\,x}_{m+1}(\mu,s,r,x,y)\Big]\,dy\,dr.\end{aligned}\tag{5.35}$$

For the first term appearing in the right-hand side of the previous identity, we use (5.11) to get rid off the singularity in time induced by $\partial_s\widehat{p}_{m+1}(\mu,s,r,x,y)$. For the second term appearing in the right-hand side of (5.35), we combine the following relation

$$\begin{aligned}\partial_s\widehat{p}^{\,z}_{m+1}(\mu,s,t,x,y)&=-\frac{1}{2}\sum_{i,j=1}^d a_{i,j}(s,z,\mu)H_2^{i,j}\left(\int_s^t a(r,z,[X_r^{s,\xi,(m)}])\,dr,y-x\right)\widehat{p}^{\,z}_{m+1}(\mu,s,t,x,y)\\&\quad+Df_{y-x}\left(\int_s^t a(r,z,[X_r^{s,\xi,(m)}])dr\right)\\&\quad\cdot\int_s^t\int\int(\widetilde{a}(r,z,y',[X_r^{s,\xi,(m)}])-\widetilde{a}(r,z,x',[X_r^{s,\xi,(m)}]))\,\partial_s p_m(\mu,s,r,x',y')\,dy'\,\mu(dx')dr\end{aligned}\tag{5.36}$$

with the η-Hölder regularity of the two maps $y\mapsto a_{i,j}(s,y,\mu),\widetilde{a}(r,z,y,\mu)$, (5.13) (at step m) and the continuity of $(s,\mu)\mapsto a(r,x,[X_r^{s,\xi,(m)}])$, $\widetilde{a}(r,x,y,[X_r^{s,\xi,(m)}])$ to deduce that $(s,x,\mu)\mapsto\partial_s\widehat{p}^{\,y}_{m+1}(\mu,s,r,x,y)-\partial_s\widehat{p}^{\,x}_{m+1}(\mu,s,r,x,y)$ is continuous and, by the space-time inequality (1.4), satisfies the inequality $|\partial_s\widehat{p}^{\,y}_{m+1}(\mu,s,r,x,y)-\partial_s\widehat{p}^{\,x}_{m+1}(\mu,s,r,x,y)|\le C(r-s)^{-1+\frac{\eta}{2}}g(c(r-s),y-x)$. From the representation formula (5.9) and the continuity of $(s,x,\mu)\mapsto\mathcal{H}_{m+1}(\mu,s,t,x,z)$, the same conclusion holds for the map $(s,x,\mu)\mapsto\Phi_{m+1}(\mu,s,t,x,z)$. We thus deduce that the second term appearing in the right-hand side of (5.35) is continuous w.r.t (s,x,μ). We then conclude that the two maps: $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\partial_s\widehat{p}_{m+1}(\mu,s,t,x,y)$, $(\partial_s\widehat{p}_{m+1}\otimes\Phi_{m+1})(\mu,s,t,x,z)$ are also continuous. Combining (5.31), (5.32) and arguments similar to those previous employed, we derive in an analogous manner that the two maps $(s,x,\mu)\mapsto p_{m+1}\otimes\partial_s\mathcal{H}_{m+1}(\mu,s,t,x,z)$, $[(p_{m+1}\otimes\partial_s\mathcal{H}_{m+1})\otimes\Phi_{m+1}](\mu,s,t,x,z)$ are continuous. We thus finally conclude that $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\partial_s p_{m+1}(\mu,s,t,x,z)$ is continuous. From the relation (5.21), the same conclusion remains valid for the maps $(s,x,\mu)\mapsto\partial_x^n p_{m+1}(\mu,s,t,x,z)$, for $n=1,2$.

We also deduce that the mappings $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto\mathcal{H}_{m+1}(\mu,s,r,t,x,z)$, $\widehat{p}^{\,y'}_{m+1}(\mu,s,r,t,x,z)$, $\widehat{p}^{\,y'}_{m+1}(\mu,s,t,x,z)$ are L-differentiable and that $\mathbb{R}^d\ni y\mapsto\partial_\mu\mathcal{H}_{m+1}(\mu,s,r,t,x,z)(y)$, $\partial_\mu\widehat{p}^{\,y'}_{m+1}(\mu,s,t,x,z)(y)$, $\partial_\mu\widehat{p}^{\,y'}_{m+1}(\mu,s,r,t,x,z)(y)$ are continuously differentiable satisfying for $0\le s\le r<t$

$$\partial_y^n[\partial_\mu\widehat{p}^{\,y'}_{m+1}(\mu,s,r,t,x,z)](y)=Df_{z-x}\Big(\int_r^t a(v,y',[X_v^{s,\xi,(m)}])\,dv\Big).\int_r^t\partial_y^n[\partial_\mu a(v,y',[X_v^{s,\xi,(m)}])](y)\,dv\tag{5.37}$$

and

$$\partial_y^n[\partial_\mu\mathcal{H}_{m+1}(\mu,s,r,t,x,z)](y)=:\mathrm{I}^n(y)+\mathrm{II}^n(y)+\mathrm{III}^n(y),\quad n=0,1,\tag{5.38}$$

with

$$\begin{aligned}
\mathrm{I}^n(y) &:= \left\{ -\sum_{i=1}^d H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x \right) \partial_y^n[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])](y) \right\} \widehat{p}_{m+1}(\mu,s,r,t,x,z) \\
&+ \left\{ -\sum_{i=1}^d b_i(r,x,[X_r^{s,\xi,(m)}]) \partial_y^n \left[\partial_\mu H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x \right) \right](y) \right\} \widehat{p}_{m+1}(\mu,s,r,t,x,z) \\
&=: \mathrm{I}_1^n(y) + \mathrm{I}_2^n(y), \\
\mathrm{II}^n(y) &:= \Big\{ \frac{1}{2} \sum_{i,j=1}^d \partial_y^n[\partial_\mu \left(a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}]) \right)](y) \\
&\quad \times H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x \right) \Big\} \widehat{p}_{m+1}(\mu,s,r,t,x,z) \\
&+ \left\{ \frac{1}{2} \sum_{i,j=1}^d \left(a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}]) \right) \right. \\
&\quad \left. \times \partial_y^n \left[\partial_\mu H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x \right) \right](y) \right\} \widehat{p}_{m+1}(\mu,s,r,t,x,z) \\
&=: \mathrm{II}_1^n(y) + \mathrm{II}_2^n(y),
\end{aligned}$$

$$\begin{aligned}
\mathrm{III}^n(y) &:= \left\{ -\sum_{i=1}^d b_i(r,x,[X_r^{s,\xi,(m)}]) H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x \right) \right. \\
&\quad \left. + \frac{1}{2} \sum_{i,j=1}^d (a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}])) H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x \right) \right\} \\
&\quad \times \partial_y^n[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,x,z)](y).
\end{aligned}$$

From the above relations and the previous continuity results, we readily derive the continuity of the mappings $(s,\mu,y) \mapsto \partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu,s,r,t,x,z)](y)$ and $(s,x,\mu,y) \mapsto \partial_y^n[\partial_\mu \widehat{p}_{m+1}^{y'}(\mu,s,r,t,x,z)](y)$. From (5.30) and (5.37), we directly get the following bounds

$$|\partial_y^n[\partial_\mu \widehat{p}_{m+1}^{y'}(\mu,s,t,x,z)](y)| \le C(t-s)^{-\frac{1+n-\eta}{2}} g(c(t-s), z-x)$$

and

$$|\partial_y^n[\partial_\mu \widehat{p}_{m+1}^{y'}(\mu,s,r,t,x,z)](y)| \le C(r-s)^{-\frac{1+n-\eta}{2}} g(c(t-r), z-x). \tag{5.39}$$

Combining the previous estimate with (5.30) and using the η-Hölder regularity of $x \mapsto a_{i,j}(r,x,\mu)$ with the space time inequality (1.4) yield

$$\begin{aligned}
|\mathrm{I}^n| &\le \frac{K}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{1+n-\eta}{2}}} g(c(t-r), z-x), \\
|\mathrm{III}^n| &\le \frac{K}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n-\eta}{2}}} g(c(t-r), z-x).
\end{aligned}$$

From the key identity (5.26), the estimates (5.6), (5.12) at step m and using the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(t,x,y,\mu)$ on the one hand or the η-Hölder regularity of $y \mapsto \widetilde{a}_{i,j}(t,x,y,\mu)$ on the other hand, we get

$$|\partial_y^n[\partial_\mu \left(a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}]) \right)](y)| \le K \left\{ \frac{|z-x|^\eta}{(r-s)^{\frac{1+n}{2}}} \wedge \frac{1}{(r-s)^{\frac{1+n-\eta}{2}}} \right\}$$

so that, by the space-time inequality (1.4), one finally obtains

$$|\mathrm{II}^n| \le K \left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}} \wedge \frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}} \right) g(c(t-r), z-x).$$

Combining the previous estimates implies

$$|\partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu,s,r,t,x,z)](y)| \leq K\left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}} \wedge \frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}}\right) g(c(t-r),z-x).$$

The previous estimates yield the (absolute) convergence of the two series $\sum_{k\geq 0}\partial_y^n[\partial_\mu \widehat{p}_{m+1}]\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(y)$ and $\sum_{k\geq 0}(p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}])\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(y)$ as well as their global continuity w.r.t the variables s,x,μ,y on $[0,r)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d$.

Formally differentiating with respect to the variables μ and then y, the relation $p_{m+1}(\mu,s,t,x,z) = \widehat{p}_{m+1}(\mu,s,t,x,z) + (p_{m+1}\otimes\mathcal{H}_{m+1})(\mu,s,t,x,z)$, we obtain

$$\begin{aligned}\partial_y^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](y) &= \partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,t,x,z)](y) + (p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}])(\mu,s,t,x,z)(y)\\ &\quad + (\partial_y^n[\partial_\mu p_{m+1}]\otimes\mathcal{H}_{m+1})(\mu,s,t,x,z)(y)\end{aligned}$$

so that a direct iteration yields the following key relation

(5.40)

$$\partial_y^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](y) = \sum_{k\geq 0}(\partial_y^n[\partial_\mu\widehat{p}_{m+1}] + p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}])\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(y),\ n=0,1.$$

Hence, $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto p_{m+1}(\mu,s,t,x,z)$ is L-differentiable and $\mathbb{R}^d\ni y\mapsto\partial_\mu p_{m+1}(\mu,s,t,x,z)(y)$ is continuously differentiable and the maps $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\ni(s,x,\mu,y)\mapsto\partial_\mu p_{m+1}(\mu,s,t,x,z)(y)$, $\partial_y[\partial_\mu p_{m+1}(\mu,s,t,x,z)](y)$ are continuous. From (5.21) and similar arguments, we also derive the continuity of the maps $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto\partial_x^n p_{m+1}(\mu,s,t,x,z)$, for $n=1,2$. We thus conclude that the mapping $[0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(s,x,\mu)\mapsto p_{m+1}(\mu,s,t,x,z)$ is in $\mathcal{C}^{1,2,2}([0,t)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$.

We now prove the estimates (5.12) to (5.19) at step $m+1$. Since their proofs are rather long, technical and relies on similar ideas and arguments, we will not prove all the announced estimates. We start with (5.12) and will deliberately omit the proofs of the estimates (5.13), (5.14) and (5.18). We introduce the following quantities for $n=0,\,1$

$$\begin{aligned}u_m^n(s,t) &:= \sup_{v\in\mathbb{R}^d}\int\mu(dx)\int(|y-x|^\eta\wedge 1)|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,y)](v)|\,dy,\\ v_m^n(s,t) &:= \sup_{v\in\mathbb{R}^d}\int\mu(dx)\int|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,y)](v)|\,dy,\end{aligned}$$

and prove by induction on m the following key inequalities:

$$u_m^n(s,t)\leq C_{m,n}(s,t)(t-s)^{-(\frac{1+n}{2}-\eta)},\qquad v_m^n(s,t)\leq C_{m,n}(s,t)(t-s)^{-\frac{(1+n-\eta)}{2}}$$

with $C_{m,n}(s,t) := \sum_{k=1}^m C^k\prod_{i=1}^k B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{(k-1)\frac{\eta}{2}}$, $C:=C(T,a,b)$ being a positive constant independent of m. This result is straightforward for $m=1$. We assume that the result holds at step m. We first remark that from (5.37) and (5.29) there exist positive constant $K:=K(T,|\widetilde{a}|_\infty,[\widetilde{a}]_H)$, $c:=c(\lambda)$, which may vary from line to line, such that for all $m\geq 1$

$$\begin{aligned}&|\partial_v^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,t,x,z)](v)|\\ &\quad\leq K\Big\{\frac{1}{(t-s)^{\frac{1+n-\eta}{2}}}+\frac{1}{t-s}\int_s^t\int\int(|y'-x'|^\eta\wedge 1)|\partial_v^n[\partial_\mu p_m(\mu,s,r,x',y')](v)|\,dy'\,\mu(dx')\,dr\Big\}\\ &\qquad\times g(c(t-s),z-x)\end{aligned}$$

so that,

$$\begin{aligned}&|\partial_v^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,t,x,z)](v)|\\ &\quad\leq K\left\{\frac{1}{(t-s)^{\frac{1+n-\eta}{2}}}+\frac{1}{(t-s)}\int_s^t\frac{C_{m,n}(s,r)}{(r-s)^{\frac{1+n}{2}-\eta}}\,dr\right\}g(c(t-s),z-x)\\ &\quad\leq K\left\{\frac{1}{(t-s)^{\frac{1+n-\eta}{2}}}+\frac{1}{(t-s)^{\frac{\eta}{2}}}\int_s^t\frac{C_{m,n}(s,r)}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}-\eta}}\,dr\right\}g(c(t-s),z-x)\\ &\quad\leq K\left\{\frac{1}{(t-s)^{\frac{1+n-\eta}{2}}}+\frac{1}{(t-s)^{\frac{1+n-\eta}{2}}}\sum_{k=1}^m C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\}\\ &\qquad\times g(c(t-s),z-x).\end{aligned}\tag{5.41}$$

Hence, by induction on r, there exists a positive constant $K := K(T, a, b, |\widetilde{a}|_\infty, [\widetilde{a}]_H)$ which may change from lines to lines but is *independent of* m such that

$$|(\partial_v^n[\partial_\mu \widehat{p}_{m+1}] \otimes \mathcal{H}_{m+1}^{(r)})(\mu, s, t, x, z)(v)| \le K^r \left\{ 1 + \sum_{k=1}^{m} C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\} \times (t-s)^{-\frac{(1+n-\eta)}{2}+r\frac{\eta}{2}} \prod_{i=1}^{r} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right) \times g(c(t-s), z-x)$$

which in turn implies

$$\begin{aligned}
&\sum_{r\ge 0} |(\partial_v[\partial_\mu \widehat{p}_{m+1}] \otimes \mathcal{H}_{m+1}^{(r)})(\mu, s, t, x, z)(v)| \\
&\quad \le \frac{K}{(t-s)^{\frac{1+n-\eta}{2}}} \left\{ 1 + \sum_{k=1}^{m} C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\} g(c(t-s), z-x) \\
&\quad \le \frac{K}{(t-s)^{\frac{1+n-\eta}{2}}} \left\{ B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2}\right) + \sum_{k=1}^{m} C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\} \\
&\qquad \times g(c(t-s), z-x). \qquad (5.42)
\end{aligned}$$

We now come back to the decomposition (5.38) of $\partial_v^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, r, t, x, z)](y)$. From (5.29) and the induction hypothesis, we directly get

$$\begin{aligned}
|\mathrm{I}_1^n(y)| &= \left| -\sum_{i=1}^{d} H_1^i \left(\int_r^t a(v, z, [X_v^{s,\xi,(m)}])dv, z-x \right) \partial_y^n \left[\partial_\mu b_i(r, x, [X_r^{s,\xi,(m)}]) \right](y) \right| \widehat{p}_{m+1}(\mu, r, t, x, z) \\
&\le K \frac{(1 + (r-s)^{\frac{1+n-\eta}{2}} u_m^n(s, r))}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{1+n-\eta}{2}}} g(c(t-r), z-x) \\
&\le K \frac{(1 + C_{m,n}(s, r)(r-s)^{\frac{\eta}{2}})}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{1+n-\eta}{2}}} g(c(t-r), z-x).
\end{aligned}$$

Next again from (5.29) and the space-time inequality (1.4) we have

$$\begin{aligned}
|\mathrm{I}_2^n(y)| &= \left| -\sum_{i=1}^{d} b_i(r, x, [X_r^{s,\xi,(m)}]) \partial_y^n \left[\partial_\mu H_1^i \left(\int_r^t a(v, z, [X_v^{s,\xi,(m)}])dv, z-x \right) \right](y) \right| \widehat{p}_{m+1}(\mu, r, t, x, z) \\
&\le K \frac{|z-x|}{(t-r)^2} \left(\int_r^t \frac{(1 + (v-s)^{\frac{1+n-\eta}{2}} u_m^n(s, v))}{(v-s)^{\frac{1+n-\eta}{2}}} dv \right) g(c(t-r), z-x) \\
&\le K \frac{(1 + (t-r)^{-1} \int_r^t C_{m,n}(s, v)(v-s)^{\frac{\eta}{2}} dv)}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{1+n-\eta}{2}}} g(c(t-r), z-x).
\end{aligned}$$

Hence, one concludes

$$\begin{aligned}
\forall y \in \mathbb{R}^d,\ |\mathrm{I}^n(y)| &\le K \frac{(1 + C_{m,n}(s, r)(r-s)^{\frac{\eta}{2}} + (t-r)^{-1} \int_r^t C_{m,n}(s, v)(v-s)^{\frac{\eta}{2}} dv)}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{1+n-\eta}{2}}} g(c(t-r), z-x) \\
&\le K \left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)} \wedge \frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}} \right) \\
&\quad \times \left(1 + C_{m,n}(s, r)(r-s)^{\frac{\eta}{2}} + (t-r)^{-1} \int_r^t C_{m,n}(s, v)(v-s)^{\frac{\eta}{2}} dv \right) g(c(t-r), z-x).
\end{aligned}$$

We now estimate II^n which is the tricky part of our computations. Since $\mu \mapsto a(r,x,\mu)$ belongs to (**CS**), $\mathrm{II}_1^n(y)$ can be written as

$$\begin{aligned}
\mathrm{II}_1^n(y) &= \frac{1}{2}\Bigg\{ \sum_{i,j=1}^d H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x\right) \\
&\quad \times \partial_y^n\Big[\partial_\nu\left[\int\{\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}])\}\, p_m(\nu,s,r,y')dy'\right]_{|\nu=\mu}\Big](y)\Bigg\} \\
&\quad \times \widehat{p}_m(\mu,r,t,x,z) \\
&=: \Big\{\frac{1}{2}\sum_{i,j=1}^d H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x\right) \times \mathrm{J}_{i,j}(y)\Big\}\,\widehat{p}_m(\mu,r,t,x,z).
\end{aligned}$$

On the one hand, using our induction hypothesis and the fact that $x \mapsto \widetilde{a}_{i,j}(r,x,y,\mu)$ is η-Hölder, one has

$$\begin{aligned}
|\mathrm{J}_{i,j}(y)| &= \left|\partial_y^n\Big[\partial_\nu\left[\int\{\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}])\}\, p_m(\nu,s,r,y')\,dy'\right]_{|\nu=\mu}\Big](y)\right| \\
&\le K\Bigg\{\int(|z-x|^\eta \wedge 1)|\partial_x^{1+n} p_m(\mu,s,r,y,y')|\,dy' \\
&\quad + \int\int(|z-x|^\eta \wedge 1)|\partial_y^n[\partial_\mu p_m(\mu,s,r,x',y')](y)|\,dy'\mu(dx')\Bigg\} \\
&\le K\frac{|z-x|^\eta}{(r-s)^{\frac{1+n}{2}}}(1+(r-s)^{\frac{1+n}{2}}v_m^n(s,r)). \qquad (5.43)
\end{aligned}$$

On the other hand, from the definition of p_m and Fubini's theorem, one gets the following decomposition

$$\begin{aligned}
\mathrm{J}_{i,j}(y) &= \partial_y^n\Big[\partial_\nu\left[\int\{\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}])\}p_m(\nu,s,r,y')\,dy'\right]_{|\nu=\mu}\Big](y) \\
&= \partial_y^n\Big[\partial_{\nu_1}\left[\int\int\{\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}])\}p_m(\nu_1,s,r,x',y')\,\mu(dx')\,dy'\right]_{|\nu_1=\mu}\Big](y) \\
&+ \partial_y^n\Big[\partial_{\nu_2}\left[\int\int\{\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}])\}p_m(\mu,s,r,x',y')\,\nu_2(dx')\,dy'\right]_{|\nu_2=\mu}\Big](y) \\
&= \Bigg[\int\int\Big\{\big(\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,x,x',[X_r^{s,\xi,(m)}])\big) \\
&\quad - \big(\widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,x',[X_r^{s,\xi,(m)}])\big)\Big\}\partial_y^n\Big[\partial_\mu p_m(\mu,s,r,x',y')\Big](y)\,\mu(dx')\,dy'\Bigg] \qquad (5.44) \\
&+ \Bigg[\int\Big\{\big(\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,x,y,[X_r^{s,\xi,(m)}])\big) \\
&\quad - \big(\widetilde{a}_{i,j}(r,z,y',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,y,[X_r^{s,\xi,(m)}])\big)\Big\}\partial_x^{1+n} p_m(\mu,s,r,y,y')\,dy'\Bigg]
\end{aligned}$$

which in turn by our induction hypothesis yields

$$|\mathrm{J}_{i,j}(y)| \quad \le \quad K\frac{(1+(r-s)^{\frac{1+n-\eta}{2}}u_m^n(s,r))}{(r-s)^{\frac{1+n-\eta}{2}}}. \qquad (5.45)$$

Consequently, combining the estimates (5.43) and (5.45), we obtain

$$|\mathrm{J}_{i,j}(y)| \le K\left\{\frac{|z-x|^\eta}{(r-s)^{\frac{1+n}{2}}} \wedge \frac{1}{(r-s)^{\frac{1+n-\eta}{2}}}\right\}(1+(r-s)^{\frac{1+n-\eta}{2}}u_m^n(s,r)+(r-s)^{\frac{1+n}{2}}v_m^n(s,r)). \qquad (5.46)$$

Hence, from the previous estimate and the space-time inequality (1.4), we deduce

$$
\begin{aligned}
|\mathrm{II}_1^n(y)| &\leq \frac{K}{t-r}\left(\frac{|z-x|^\eta}{(r-s)^{\frac{1+n}{2}}}\wedge\frac{1}{(r-s)^{\frac{1+n-\eta}{2}}}\right)(1+C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}})\,g(c(t-r),z-x)\\
&\leq K\left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}}\right)(1+C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}})\,g(c(t-r),z-x).
\end{aligned}
$$

We now turn to estimate II_2^n. From the very definition of $H_2^{i,j}$, (5.29) and noticing that for any differentiable map $\mathcal{P}_2(\mathbb{R}^d)\ni\nu\mapsto\Sigma(\nu)$ taking values in the set of positive definite matrix one has $\partial_y[\partial_\mu(\Sigma^{-1}(\mu))_{i,j}](y)=-(\Sigma^{-1}(\mu)\partial_y[\partial_\mu\Sigma(\mu)](y)\Sigma^{-1}(\mu))_{i,j}=-\partial_y\partial_\nu[\sum_{k_1,k_2}(\Sigma^{-1}(\mu))_{i,k_1}(\Sigma(\nu))_{k_1,k_2}(\Sigma^{-1}(\mu))_{k_2,j}]_{\nu=\mu}(y)$, we get

$$
\begin{aligned}
&\left|\partial_y^n\partial_\mu\left[H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv,z-x\right)\right](y)\right|\\
&\quad\leq K\left(\frac{|z-x|^2}{(t-r)^3}+\frac{1}{(t-r)^2}\right)\max_{i,j}\left|\int_r^t\partial_y^n\partial_\mu[a_{i,j}(v,z,[X_v^{s,\xi,(m)}])](y)dv\right|\\
&\quad\leq K\left(\frac{|z-x|^2}{(t-r)^2}+\frac{1}{t-r}\right)\frac{(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)}{(r-s)^{\frac{1+n-\eta}{2}}}
\end{aligned}
$$

so that, from the space-time inequality (1.4), we clearly deduce

$$
\begin{aligned}
|\mathrm{II}_2^n(y)| &\leq K\left(\frac{|z-x|^{2+\eta}}{(t-r)^2}+\frac{|z-x|^\eta}{(t-r)}\right)\frac{(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)}{(r-s)^{\frac{1+n-\eta}{2}}}\,g(c(t-r),z-x)\\
&\leq\frac{K}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n-\eta}{2}}}(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)\,g(c(t-r),z-x)\\
&\leq K\left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}}\right)(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)\\
&\quad\times g(c(t-r),z-x).
\end{aligned}
$$

Hence, gathering estimates on II_1^n and II_2^n, we get for all $y\in\mathbb{R}^d$

$$
\begin{aligned}
|\mathrm{II^n(y)}| &\leq K\left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}}\right)(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)\\
&\quad\times g(c(t-r),z-x).
\end{aligned}
$$

Finally, for $r\neq s$, from the relation (5.37) and the estimate (5.29), we get

$$
\left|\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y)\right|\leq\frac{K}{(r-s)^{\frac{1+n-\eta}{2}}}(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)\,g(c(t-r),z-x)
$$

so that

$$
\begin{aligned}
|\mathrm{III}^n(y)| &\leq K\left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}}\right)(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)\\
&\quad\times g(c(t-r),z-x).
\end{aligned}
$$

Gathering the previous estimates together, we finally obtain

$$
\begin{aligned}
|\partial_y^n[\partial_\mu\mathcal{H}_{m+1}(\mu,s,r,t,x,z)](y)| &\leq K\left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s)^{\frac{1+n-\eta}{2}}}\right)\\
&\quad\times\left[1+C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}}+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv\right]\\
&\quad\times g(c(t-r),z-x).
\end{aligned}
\tag{5.47}
$$

Now, our aim is to establish an upper-bound of the quantity $p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}](\mu,r,t,x,z)$. The estimate (5.47) allows to balance the singularity in time induced by $\partial_y^n[\partial_\mu\mathcal{H}_{m+1}]$. Indeed, assuming first

that $r \in [s, \frac{t+s}{2}]$, one has $t - r \geq (t-s)/2$ which directly implies

$$\begin{aligned}\int |p_{m+1}(\mu, s, r, x, y')||\partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, r, t, y', z)](y)|dy' \\ \leq \frac{K}{(t-s)(r-s)^{\frac{1+n-\eta}{2}}}\left(1 + C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}} + (t-s)^{-1}\int_s^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv\right) \\ \times g(c(t-s), z-x)\end{aligned}$$

so that

$$\begin{aligned}&\int_s^{\frac{t+s}{2}} \int |p_{m+1}(\mu, s, r, x, y')||\partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, r, t, y', z)](y)|dy'\,dr \\ &\quad \leq K\left\{\frac{B(1, \frac{1-n+\eta}{2})}{(t-s)^{\frac{1+n-\eta}{2}}} + \frac{1}{t-s}\int_s^t \frac{C_{m,n}(s,r)}{(r-s)^{\frac{1+n}{2}-\eta}}\,dr\right\} g(c(t-s), z-x) \\ &\quad \leq K\left\{\frac{B(\frac{\eta}{2}, \frac{1-n+\eta}{2})}{(t-s)^{\frac{1+n-\eta}{2}}} + \frac{1}{(t-s)^{\frac{\eta}{2}}}\int_s^t \frac{C_{m,n}(s,r)}{(t-s)^{1-\frac{\eta}{2}}(r-s)^{\frac{1+n}{2}-\eta}}\,dr\right\} \\ &\qquad \times g(c(t-s), z-x) \\ &\quad \leq \frac{K}{(t-s)^{\frac{1+n-\eta}{2}}}\left\{B\left(\frac{\eta}{2}, \frac{1+\eta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\} \\ &\qquad \times g(c(t-s), z-x).\end{aligned}$$

Then, assuming that $r \in [\frac{t+s}{2}, t]$, one has $r - s \geq (t-s)/2$ so that

$$\begin{aligned}\int |p_{m+1}(\mu, s, r, x, y')||\partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, r, t, y', z)](y)|dy' \\ \leq \frac{K}{(t-s)^{\frac{1+n}{2}}(t-r)^{1-\frac{\eta}{2}}}(1 + C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}} + (t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv) \\ \times g(c(t-s), z-x)\end{aligned}$$

which in turn, by Fubini's theorem, directly yields

$$\begin{aligned}&\int_{\frac{t+s}{2}}^t \int |p_{m+1}(\mu, s, r, x, y')||\partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, r, t, y', z)](y)|\,dy'\,dr \\ &\quad \leq \frac{K}{(t-s)^{\frac{1+n}{2}}}\int_{\frac{t+s}{2}}^t \frac{1}{(t-r)^{1-\frac{\eta}{2}}}(1 + C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}} + (t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv)\,dr \\ &\qquad \times g(c(t-s), z-x) \\ &\quad \leq K\left\{\frac{B(1, \frac{\eta}{2})}{(t-s)^{\frac{1+n-\eta}{2}}} + \frac{1}{(t-s)^{\frac{\eta}{2}}}\int_s^t \frac{C_{m,n}(s,v)}{(t-v)^{1-\frac{\eta}{2}}(v-s)^{\frac{1+n}{2}-\eta}}\,dv\right\} g(c(t-s), z-x) \\ &\quad \leq \frac{K}{(t-s)^{\frac{1+n-\eta}{2}}}\left\{B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\} \\ &\qquad \times g(c(t-s), z-x).\end{aligned}$$

Gathering the two previous cases, we clearly obtain

$$\begin{aligned}&|p_{m+1} \otimes \partial_y^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, t, x, z)](y)| \\ &\leq \frac{K}{(t-s)^{\frac{1+n-\eta}{2}}}\left(B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} + (i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right) \\ &\quad \times g(c(t-s), z-x)\end{aligned} \tag{5.48}$$

so that

$$\sum_{r\geq 0}|(p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}])\otimes\mathcal{H}_{m+1}^{(r)}(\mu,s,t,x,z)(y)|$$
$$\leq\frac{K}{(t-s)^{\frac{1+n-\eta}{2}}}\left\{B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}\right)+\sum_{k=1}^{m}C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\}$$
$$\times g(c(t-s),z-x). \tag{5.49}$$

The estimates (5.42) and (5.49) together with the representation formula (5.40) imply that there exist two constants K, c such that

$$|\partial_y^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](y)|$$
$$\leq\frac{K}{(t-s)^{\frac{1+n-\eta}{2}}}\left\{B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}\right)+\sum_{k=1}^{m}C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\}$$
$$\times g(c(t-s),z-x)$$

so that

$$u_{m+1}^n(s,t)\leq\frac{K}{(t-s)^{\frac{1+n}{2}-\eta}}\left\{B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}\right)+\sum_{k=1}^{m}C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\}$$

and similarly,

$$v_{m+1}^n(s,t)\leq\frac{K}{(t-s)^{\frac{1+n-\eta}{2}}}\left\{B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}\right)+\sum_{k=1}^{m}C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\}.$$

Since the constant K does not depend either on the constant C appearing in the definition of $C_{m,n}(s,t)$ or m, one may change C once for all and derive the induction hypothesis at step $m+1$ for u_{m+1}^n and v_{m+1}^n. This completes the proof of (5.12).

The estimate (5.13) at step $m+1$ follows by combining the relations (5.34), (5.31), (5.36) with estimates analogous to the one established above while the estimate (5.14) follows from the representation formula (5.40). The remaining technical details are omitted.

In order to derive (5.15) at step $m+1$, we proceed similarly. We introduce the quantities

$$u_m(s,t)$$
$$:=\sup_{(y,y')\in(\mathbb{R}^d)^2,\,y\neq y'}\int\mu(dx')\int(|y''-x'|^\eta\wedge 1)\frac{|\partial_y[\partial_\mu p_m(\mu,s,t,x,y'')](y)-|\partial_y[\partial_\mu p_m(\mu,s,t,x,y'')](y')|}{|y-y'|^\beta}\,dy'',$$
$$v_m(s,t)$$
$$:=\sup_{(y,y')\in(\mathbb{R}^d)^2,\,y\neq y'}\int\mu(dx')\int\frac{|\partial_y[\partial_\mu p_m(\mu,s,t,x,y'')](y)-|\partial_y[\partial_\mu p_m(\mu,s,t,x,y'')](y')|}{|y-y'|^\beta}\,dy'',$$

for any $\beta\in[0,\eta)$ and $m\geq 1$. We prove by induction the following key inequalities:

$$u_m(s,t)\leq C_m(s,t)(t-s)^{-(1+\frac{\beta}{2}-\eta)},\quad\text{and}\quad v_m(s,t)\leq C_m(s,t)(t-s)^{-(1+\frac{\beta-\eta}{2})},$$

with $C_m(s,t):=\sum_{k=1}^m C^k\prod_{i=1}^k B\left(\frac{\eta}{2},\frac{\eta-\beta}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{(k-1)\frac{\eta}{2}}$, $C:=C(T,a,b,|\widetilde{a}|_\infty,[\widetilde{a}]_H,|\widetilde{b}|_\infty,[\widetilde{b}]_H)$ being a positive constant independent of m. The result being straightforward for $m=1$, we assume that it holds at step m. By direct computations, we first remark that the following decomposition

$$\partial_y[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y)-\partial_y[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y')=Df_{z-x}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv\right).$$

(5.50)

$$\int_r^t\left\{\int[\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}])-\widetilde{a}(v,z,y_0,[X_v^{s,\xi,(m)}])][\partial_x^2 p_m(\mu,s,v,y,y'')-\partial_x^2 p_m(\mu,s,v,y',y'')]dy''\right.$$
$$+\int\int[\widetilde{a}(v,z,y'',[X_v^{s,\xi,(m)}])-\widetilde{a}(v,z,x',[X_v^{s,\xi,(m)}])]$$
$$.\Big(\partial_y[\partial_\mu p_m(\mu,s,v,x',y'')](y)-\partial_y[\partial_\mu p_m(\mu,s,v,x',y'')](y')\Big)\,dy''\,\mu(dx')\Big\}\,dv$$

holds for any $y_0\in\mathbb{R}^d$.

We split the computations into the two following cases: $|y-y'|^2 \le v-s$ and $|y-y'|^2 \ge v-s$. In the first case, we choose $y_0 = y$. From (5.7) with $n=2$, $\beta \in [0,\eta)$ and the space-time inequality (1.4), we get

$$\begin{aligned}
&\Big| \int [\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,y_0,[X_r^{s,\xi,(m)}])][\partial_x^2 p_m(\mu,s,v,y,y'') - \partial_x^2 p_m(\mu,s,v,y',y'')]dy''\Big| \\
&\quad \le K|y-y'|^\beta \int \frac{|y''-y|^\eta}{(v-s)^{1+\frac{\beta}{2}}} \left\{ g(c(v-s),y''-y) + g(c(v-s),y''-y') \right\} dy'' \\
&\quad \le K|y-y'|^\beta \int \frac{1}{(v-s)^{1+\frac{\beta}{2}}} \Big\{ |y''-y|^\eta g(c(v-s),y''-y) + (|y''-y'|^\eta + |v-s|^{\frac{\eta}{2}}) g(c(v-s),y''-y') \Big\} dy'' \\
&\quad \le K \frac{|y-y'|^\beta}{(v-s)^{1+\frac{\beta-\eta}{2}}}.
\end{aligned} \tag{5.51}$$

Otherwise, if $|y-y'|^2 \ge v-s$ then from the Hölder regularity of $y' \mapsto \widetilde{a}(t,z,y',\mu)$, (5.6) for $n=2$ and the space-time inequality (1.4), we get the same inequality. Plugging the previous bound in (5.50) yields

$$\begin{aligned}
|\partial_y[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,x,z)](y) - \partial_y[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,x,z)](y')| &\le K \frac{|y-y'|^\beta}{t-r} \Big\{ \int_r^t \Big[\frac{1}{(v-s)^{1+\frac{(\beta-\eta)}{2}}} \\
&+ \int\int (|y''-x'|^\eta \wedge 1) \frac{|\partial_v[\partial_\mu p_m(\mu,s,v,x',y'')](y) - \partial_v[\partial_\mu p_m(\mu,s,v,x',y'')](y')|}{|y-y'|^\beta} dy'' \, \mu(dx') \Big] dv \Big\} \\
&\times g(c(t-r),z-x)
\end{aligned} \tag{5.52}$$

for any $\beta \in [0,\eta)$ so that, from the previous inequality with $r=s$ and the induction hypothesis,

$$\begin{aligned}
&|\partial_y[\partial_\mu \widehat{p}_{m+1}(\mu,s,t,x,z)](y) - \partial_y[\partial_\mu \widehat{p}_{m+1}(\mu,s,t,x,z)](y')| \\
&\quad \le K|y-y'|^\beta \left\{ \frac{1}{(t-s)^{1-\frac{(\eta-\beta)}{2}}} + \frac{1}{t-s} \int_s^t \frac{C_m(s,r)}{(r-s)^{1+\frac{\beta}{2}-\eta}} dr \right\} g(c(t-s),z-x) \\
&\quad \le K|y-y'|^\beta \left\{ \frac{1}{(t-s)^{1-\frac{(\eta-\beta)}{2}}} + \frac{1}{(t-s)^{\frac{\eta}{2}}} \int_s^t \frac{C_m(s,r)}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\beta}{2}-\eta}} dr \right\} g(c(t-s),z-x) \\
&\quad = K \frac{|y-y'|^\beta}{(t-s)^{1-\frac{(\eta-\beta)}{2}}} \left\{ 1 + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\} g(c(t-s),z-x).
\end{aligned} \tag{5.53}$$

Introducing the notation $\Delta_y \partial_y[\partial_\mu \widehat{p}_m(\mu,s,t,x,z)](y) := \partial_y[\partial_\mu \widehat{p}_m(\mu,s,t,x,z)](y) - \partial_y[\partial_\mu \hat{p}_m(\mu,s,t,x,z)](y')$, by induction on r, there exists a positive constant $K := K(T,a,b)$ (which may change from lines to lines but is independent of m and C) such that

$$\begin{aligned}
&|(\Delta_y \partial_y \partial_\mu \widehat{p}_{m+1} \otimes \mathcal{H}_{m+1}^{(r)})(\mu,s,t,x,z)(y)| \\
&\quad \le K^r |y-y'|^\beta \left\{ 1 + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\} (t-s)^{-1+\frac{(\eta-\beta)}{2}+r\frac{\eta}{2}} \\
&\qquad \times \prod_{i=1}^r B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) g(c(t-s),z-x)
\end{aligned}$$

which in turn implies

$$\begin{aligned}
&\sum_{r\ge 0} |(\Delta_y \partial_y \partial_\mu \widehat{p}_{m+1} \otimes \mathcal{H}_{m+1}^{(r)})(\mu,s,t,x,z)(y)| \\
&\quad \le K \frac{|y-y'|^\beta}{(t-s)^{1-\frac{(\eta-\beta)}{2}}} \left\{ 1 + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\} \\
&\qquad \times g(c(t-s),z-x) \\
&\quad \le K \frac{|y-y'|^\beta}{(t-s)^{1-\frac{(\eta-\beta)}{2}}} \left\{ B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\} \\
&\qquad \times g(c(t-s),z-x).
\end{aligned} \tag{5.54}$$

Then, from (5.38), we write

$$\begin{aligned}\partial_y[\partial_\mu \mathcal{H}_m(\mu,s,r,t,x,z)](y) - \partial_y[\partial_\mu \mathcal{H}_m(\mu,s,r,t,x,z)](y') \\ &= \mathrm{I}^1(y) - \mathrm{I}^1(y') + \mathrm{II}^1(y) - \mathrm{II}^1(y') + \mathrm{III}^1(y) - \mathrm{III}^1(y') \\ &= \mathrm{I}_1^1(y) - \mathrm{I}_1^1(y') + \mathrm{I}_2^1(y) - \mathrm{I}_2^1(y') + \mathrm{II}_1^1(y) - \mathrm{II}_1^1(y') + \mathrm{II}_2^1(y) - \mathrm{II}_2^1(y') \\ &\quad + \mathrm{III}^1(y) - \mathrm{III}^1(y')\end{aligned}$$

and prove appropriate estimates for each terms. With our notations, one has

$$\begin{aligned}\mathrm{I}_1^1(y) - \mathrm{I}_1^1(y') &= \sum_{i=1}^d H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x\right) \\ &\quad \times \left\{\partial_y\left[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])\right](y) - \partial_y\left[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])\right](y')\right\}\widehat{p}_{m+1}(\mu,r,t,x,z).\end{aligned}$$

Moreover, from (5.27), one gets

$$\begin{aligned}&\partial_y\left[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])\right](y) - \partial_y\left[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])\right](y') \\ &= \int (\widetilde{b}_i(r,x,z,[X_r^{s,\xi,(m)}]) - \widetilde{b}_i(r,x,v_0,[X_r^{s,\xi,(m)}])).(\partial_x^2 p_m(\mu,s,r,y,z) - \partial_x^2 p_m(\mu,s,r,y',z))dz \\ &\quad + \int\int (\widetilde{b}_i(r,x,z,[X_r^{s,\xi,(m)}]) - \widetilde{b}_i(r,x,x',[X_r^{s,\xi,(m)}])) \\ &\quad .(\partial_v[\partial_\mu p_m(\mu,s,t,x',z)](y) - \partial_v[\partial_\mu p_m(\mu,s,t,x',z)](y'))\,dz\,\mu(dx')\end{aligned} \tag{5.55}$$

which is valid for any $v_0 \in \mathbb{R}^d$. Hence, using (5.7) and the uniform η-Hölder regularity of $z \mapsto \widetilde{b}_i(r,x,z,\mu)$, as well as similar arguments as those employed in order to establish (5.51) for the first part appearing in the right-hand side of the previous equality, we get

$$\begin{aligned}\forall \beta \in [0,\eta), \quad &|\partial_y\left[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])\right](y) - \partial_y\left[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])\right](y')| \\ &\leq K\frac{|y-y'|^\beta}{(r-s)^{1+\frac{(\beta-\eta)}{2}}}\left[1 + (r-s)^{1+\frac{(\beta-\eta)}{2}}u_m(s,r)\right]\end{aligned} \tag{5.56}$$

which in turn implies

$$\begin{aligned}|\mathrm{I}_1^1(y) - \mathrm{I}_1^1(y')| &\leq K\frac{|y-y'|^\beta}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}}\left[1 + (r-s)^{1+\frac{(\beta-\eta)}{2}}u_m(s,r)\right]\, g(c(t-r),z-x) \\ &\leq K\frac{|y-y'|^\beta}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}}\left[1 + C_m(s,r)(r-s)^{\frac{\eta}{2}}\right]\, g(c(t-r),z-x).\end{aligned}$$

Next, we proceed similarly

$$\begin{aligned}|\mathrm{I}_2^1(y) - \mathrm{I}_2^1(y')| &:= \Big|\sum_{i=1}^d b_i(r,x,[X_r^{s,\xi,(m)}])\left\{\partial_y\left[\partial_\mu H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x\right)\right](y)\right. \\ &\quad \left. -\partial_y\left[\partial_\mu H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x\right)\right](y')\right\}\Big|\,\widehat{p}_m(\mu,r,t,x,z) \\ &\leq \frac{K}{(t-r)^{\frac{3}{2}}}\int_r^t \max_{i,j}\Big|\partial_y[\partial_\mu a_{i,j}(v,z,[X_v^{s,\xi,(m)}])](y) - \partial_y[\partial_\mu a_{i,j}(v,z,[X_v^{s,\xi,(m)}])](y')\Big|\,dv \\ &\quad \times g(c(t-r),z-x).\end{aligned}$$

From (5.26) and similar arguments as those employed to establish (5.56), that is, the decomposition (5.55) with the maps $a_{i,j}$ and $\widetilde{a}_{i,j}$ instead of b_i and $\widetilde{b}_i$, in a completely analogous manner, one gets

$$\begin{aligned}&\Big|\partial_y\left[\partial_\mu a_{i,j}(v,z,[X_v^{s,\xi,(m)}])\right](y) - \partial_y\left[\partial_\mu a_{i,j}(v,z,[X_v^{s,\xi,(m)}])\right](y')\Big| \\ &\qquad \leq K\frac{|y-y'|^\beta}{(v-s)^{1+\frac{(\beta-\eta)}{2}}}(1 + C_m(s,v)(v-s)^{\frac{\eta}{2}})\end{aligned} \tag{5.57}$$

for any $\beta \in [0,\eta)$, so that we obtain

$$|\mathrm{I}_2^1(y) - \mathrm{I}_2^1(y')| \leq K\frac{|y-y'|^\beta}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}}\Big(1 + (t-r)^{-1}\int_r^t C_m(s,v)(v-s)^{\frac{\eta}{2}}\,dv\Big)\, g(c(t-r),z-x).$$

Gathering the two previous estimates, we conclude

$$|\mathrm{I}^1(y)-\mathrm{I}^1(y')| \leq K\frac{|y-y'|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}}\Big(1+C_m(s,r)(r-s)^{\frac{\eta}{2}}+(t-r)^{-1}\int_r^t C_m(s,v)(v-s)^{\frac{\eta}{2}}\Big)\, g(c(t-r),z-x).$$

Still using our notations, we have

$$\mathrm{II}_1^1(y)-\mathrm{II}_1^1(y') \;=\; \sum_{i,j=1}^d \Big\{H_2^{i,j}\Big(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-x\Big)\times(\mathrm{J}_{i,j}(y)-\mathrm{J}_{i,j}(y'))\Big\}\,\widehat{p}_{m+1}(\mu,r,t,x,z).$$

On the one hand, using (5.44) (with $n=1$), the induction hypothesis and the fact that $x\mapsto \widetilde{a}_{i,j}(r,x,y,\mu)$ is η-Hölder uniformly with respect to the other variables, we obtain

$$|\mathrm{J}_{i,j}(y)-\mathrm{J}_{i,j}(y')| \;\leq\; K\frac{|z-x|^\eta|y-y'|^\beta}{(r-s)^{1+\frac{\beta}{2}}}(1+(r-s)^{1+\frac{\beta}{2}}v_m(s,r)).$$

On the other hand, similarly to (5.44), one gets the following decomposition

$$\begin{aligned}
\mathrm{J}_{i,j}(y)-\mathrm{J}_{i,j}(y') = \int\int \Big\{&\big(\widetilde{a}_{i,j}(r,x,y'',[X_r^{s,\xi,(m)}])-\widetilde{a}_{i,j}(r,x,x',[X_r^{s,\xi,(m)}])\big)\\
&-\Big(\widetilde{a}_{i,j}(r,z,y'',[X_r^{s,\xi,(m)}])-\widetilde{a}_{i,j}(r,z,x',[X_r^{s,\xi,(m)}])\Big)\Big\}\\
&.\Big[\partial_y[\partial_\mu p_m(\mu,s,r,x',y'')](y)-\partial_y[\partial_\mu p_m(\mu,s,r,x',y'')](y')\Big]\,dy''\,\mu(dx')\\
+\int\Big\{&\big(\widetilde{a}_{i,j}(r,x,y'',[X_r^{s,\xi,(m)}])-\widetilde{a}_{i,j}(r,x,v_0,[X_r^{s,\xi,(m)}])\big)\\
&-\big(\widetilde{a}_{i,j}(r,z,y'',[X_r^{s,\xi,(m)}])-\widetilde{a}_{i,j}(r,z,v_0,[X_r^{s,\xi,(m)}])\big)\Big\}\\
&.[\partial_x^2 p_m(\mu,s,r,y,y'')-\partial_x^2 p_m(\mu,s,r,y',y'')]\,dy''
\end{aligned}$$

which is valid for any $v_0\in\mathbb{R}^d$. Now, using the fact that $\mu\mapsto a(r,x,\mu)$ belongs to $(\mathbf{CS}_+)$ namely the fact that $y\mapsto\widetilde{a}_{i,j}(t,z,y,\mu)$ is η-Hölder, we get that the first term appearing in the right-hand side of the above decomposition is bounded by $K|y-y'|^\beta u_m(s,r)$. For the second term, one has to consider the two disjoint cases: $|y-y'|\leq(r-s)^{\frac12}$ and $|y-y'|>(r-s)^{\frac12}$. In the first case, one selects $v_0=y$ and uses the estimate (5.7) (with $n=2$), the η-Hölder regularity of $y''\mapsto\widetilde{a}_{i,j}(r,x,y'',\mu)$ as well as the inequality $|y''-y|^\eta\leq|y''-y'|^\eta+(r-s)^{\frac{\eta}{2}}$. In the second case, that is, $|y-y'|\geq(r-s)^{\frac12}$, one directly uses (5.6). To be more specific, one decompose the second term as the sum

$$\begin{aligned}
&\int[\widetilde{a}_{i,j}(r,x,y'',[X_r^{s,\xi,(m)}])-\widetilde{a}_{i,j}(r,x,y,[X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,y,y'')\,dy''\\
&\quad+\int[\widetilde{a}_{i,j}(r,x,y'',[X_r^{s,\xi,(m)}])-\widetilde{a}_{i,j}(r,x,y',[X_r^{s,\xi,(m)}])]\partial_x^2 p_m(\mu,s,r,y',y'')\,dy''
\end{aligned}$$

and bound each term using (5.6), the η-Hölder regularity of $y\mapsto\widetilde{a}_{i,j}(r,x,y,\mu)$ as well as the space-time inequality (1.4). In both cases, one concludes that the second term is bounded by $K|y-y'|^\beta(r-s)^{-1-\frac{\beta-\eta}{2}}$. We thus finally obtain

$$|\mathrm{J}_{i,j}(y)-\mathrm{J}_{i,j}(y')|\leq K\frac{|y-y'|^\beta}{(r-s)^{1+\frac{(\beta-\eta)}{2}}}(1+(r-s)^{1+\frac{(\beta-\eta)}{2}}u_m(s,r)).$$

Combining the two previous estimates yields

$$\begin{aligned}
\forall(y,y')\in(\mathbb{R}^d)^2, |\mathrm{II}_1^1(y)-\mathrm{II}_1^1(y')| \leq K&\left\{\frac{|y-y'|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\beta}{2}}}\wedge\frac{|y-y'|^\beta}{(t-r)(r-s)^{1+\frac{(\beta-\eta)}{2}}}\right\}\\
&\times\Big[1+C_m(s,r)(r-s)^{\frac{\eta}{2}}\Big]\,g(c(t-r),z-x)
\end{aligned}$$

for any $\beta\in[0,\eta)$.

Using (5.57), for all $\beta\in[0,\eta)$ and for all $(y,y')\in(\mathbb{R}^d)^2$, one gets

$$\begin{aligned}
|\mathrm{II}_2^1(y)-\mathrm{II}_2^1(y')| &\leq K\frac{|z-x|^{2+\eta}}{(t-r)^3}\int_r^t\max_{i,j}\Big|\partial_y[\partial_\mu a_{i,j}(v,z,[X_v^{s,\xi,(m)}])](y)-\partial_y[\partial_\mu a_{i,j}(v,z,[X_v^{s,\xi,(m)}])](y')\Big|\,dv\\
&\quad\times g(c(t-r),z-x)\\
&\leq K\frac{|y-y'|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}}\Big[1+(t-r)^{-1}\int_r^t C_m(s,v)(v-s)^{\frac{\eta}{2}}\,dv\Big]\,g(c(t-r),z-x).
\end{aligned}$$

Finally, for the last term, from (5.52) and the induction hypothesis, we obtain

$$\begin{aligned}\forall \beta \in [0,\eta),\ \forall (y,y') \in (\mathbb{R}^d)^2, \quad & |\mathrm{III}^1(y) - \mathrm{III}^1(y')| \\ & \leq K \frac{1}{(t-r)^{1-\frac{\eta}{2}}} |\partial_y[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,x,z)](y) - \partial_y[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,x,z)](y')| \\ & \leq K \frac{|y-y'|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}} \Big[1 + \int_r^t C_m(s,v)(v-s)^{\frac{\eta}{2}}\, dv\Big]\, g(c(t-r), z-x).\end{aligned}$$

Gathering all the previous computations, we finally conclude

$$\begin{aligned}\Big|\Delta_y \partial_y \Big[\partial_\mu \mathcal{H}_{m+1}(\mu,s,r,t,x,z)\Big](y)\Big| \leq K & \left\{ \frac{|y-y'|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\beta}{2}}} \wedge \frac{|y-y'|^\beta}{(t-r)(r-s)^{1-\frac{(\eta-\beta)}{2}}} \right\} \\ & \times \Big[1 + C_m(s,r)(r-s)^{\frac{\eta}{2}} + (t-r)^{-1} \int_r^t C_m(s,v)(v-s)^{\frac{\eta}{2}}\, dv\Big] \\ & \times g(c(t-r), z-x)\end{aligned}$$

for any $(y,y') \in (\mathbb{R}^d)^2$ and for any $\beta \in [0,\eta)$. We again separate the space-time convolution into the two disjoint cases: $r \in [s, \frac{t+s}{2}]$ and $r \in [\frac{t+s}{2}, t]$. Skipping technical details, we obtain

$$\begin{aligned}& |p_{m+1} \otimes \Delta_y \partial_y [\partial_\mu \mathcal{H}_{m+1}(\mu,s,t,x,z)](y)| \\ & \quad \leq K|y-y'|^\beta \left(\frac{1}{(t-s)^{1+\frac{(\beta-\eta)}{2}}} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) + \frac{1}{(t-s)^{\frac{\eta}{2}}} \int_s^t \frac{C_m(s,r)(r-s)^{\frac{\eta}{2}}}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{(\beta-\eta)}{2}}}\, dr \right) \\ & \qquad \times g(c(t-s), z-x) \\ & \quad \leq K \frac{|y-y'|^\beta}{(t-s)^{1+\frac{(\beta-\eta)}{2}}} \left(B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right) \\ & \qquad \times g(c(t-s), z-x)\end{aligned}$$

so that

$$\begin{aligned}& \sum_{r\geq 0} |(p_{m+1} \otimes \Delta_y \partial_y \partial_\mu \mathcal{H}_{m+1}) \otimes \mathcal{H}^{(r)}_{m+1}(\mu,s,t,x,z)(y)| \\ & \qquad \leq K \frac{|y-y'|^\beta}{(t-s)^{1+\frac{(\beta-\eta)}{2}}} \Big(B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) \\ & \qquad\quad + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \Big) g(c(t-s), z-x).\end{aligned} \tag{5.58}$$

Now, combining (5.54) and (5.58) with the following representation

$$\Delta_y \partial_y [\partial_\mu p_{m+1}(\mu,s,t,x,z)](y) = \sum_{r\geq 0} [\Delta_y \partial_y \partial_\mu \widehat{p}_{m+1} + p_{m+1} \otimes \Delta_y \partial_y \partial_\mu \mathcal{H}_{m+1}] \otimes \mathcal{H}^{(r)}_{m+1}(\mu,s,t,x,z)(y)$$

we deduce that there exist two constants K, c (independent of C and m) such that

$$\begin{aligned}|\Delta_y \partial_y [\partial_\mu p_{m+1}(\mu,s,t,x,z)](y)| \leq K & \frac{|y-y'|^\beta}{(t-s)^{1+\frac{(\beta-\eta)}{2}}} \\ & \times \left\{ B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\} \\ & \times g(c(t-s), z-x)\end{aligned}$$

so that

$$u_{m+1}(s,t) \leq \frac{K}{(t-s)^{1+\frac{\beta}{2}-\eta}} \left\{ B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\}$$

and similarly,

$$v_{m+1}(s,t) \leq \frac{K}{(t-s)^{1+\frac{\beta-\eta}{2}}} \left\{ B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2}, \frac{\eta-\beta}{2} + (i-1)\frac{\eta}{2}\right) (t-s)^{k\frac{\eta}{2}} \right\}.$$

Since the constant K does not depend either on the constant C appearing in the definition of $C_m(s,t)$ or m, one may change C once for all and derive the induction hypothesis at step m for $u_{m+1}(s,t)$ and $v_{m+1}(s,t)$. This completes the proof of (5.15).

We now establish the estimates (5.16) and (5.17). We first prove that there exist some positive constants $K := K(T,a,b)$, $c := c(\lambda)$ such that for all $m \geq 1$, for all $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$

$$\forall \alpha \in [0,1],\ |p_m(\mu,s,t,x,z) - p_m(\mu',s,t,x,z)| \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{\alpha-\eta}{2}}}\, g(c(t-s), z-x). \tag{5.59}$$

We first remark that if $W_2(\mu,\mu') \leq (t-s)^{1/2}$, then for any $\xi, \xi' \in \mathbb{L}^2$ such that $[\xi] = \mu$ and $[\xi'] = \mu'$, one has

$$\begin{aligned} p_m(\mu,s,t,x,z) - p_m(\mu',s,t,x,z) &= \int_0^1 \partial_\lambda p_m(\mu' + \lambda(\mu-\mu'), s, r, x, z)\, d\lambda \\ &= \int_0^1 \mathbb{E}\Big[\partial_\mu p_m(\nu,s,r,x,z)_{|\nu = \mu' + \lambda(\mu-\mu')}(\xi + \lambda(\xi-\xi'))(\xi-\xi')\Big] d\lambda \end{aligned}$$

so that, using (5.12) and optimising over joint distributions with μ as first marginal and μ' as second marginal

$$|p_m(\mu,s,r,x,z) - p_m(\mu',s,r,x,z)| \leq K \frac{W_2(\mu,\mu')}{(t-s)^{\frac{1-\eta}{2}}} g(c(t-s), z-x) \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{\alpha-\eta}{2}}} g(c(t-s), z-x).$$

Otherwise, if $W_2(\mu,\mu') > (t-s)^{1/2}$, we directly get

$$\forall m \geq 1,\ |p_m(\mu,s,r,x,z) - p_m(\mu',s,r,x,z)| \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{\alpha}{2}}} g(c(t-s), z-x)$$

so that, from (**HR**)(ii), for any $m \geq 1$ and any $\alpha \in [0,1]$

$$\begin{aligned} &|a_{i,j}(r,z,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi',(m)}])| \\ &\qquad \leq \Big| \int\int A_{i,j}(r,z,y',[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}]) p_m(\mu,s,r,x',y')\, dy'\, (\mu-\mu')(dx') \Big| \\ &\qquad + \Big| \int\int (A_{i,j}(r,z,y',[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}]) - A_{i,j}(r,z,x',[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}])) \\ &\qquad\quad \times (p_m(\mu,s,r,x',y') - p_m(\mu',s,r,x',y'))\, dy' \mu'(dx') \Big| \\ &\qquad \leq K \left\{ \frac{W_2^\alpha(\mu,\mu')}{(r-s)^{\frac{\alpha-\eta}{2}}} + \int (|y'-x'|^\eta \wedge 1) |p_m(\mu,s,r,x',y') - p_m(\mu',s,r,x',y')|\, dy' \mu'(dx') \right\} \\ &\qquad \leq K \frac{W_2^\alpha(\mu,\mu')}{(r-s)^{\frac{\alpha-\eta}{2}}} \end{aligned} \tag{5.60}$$

where we used the fact that $x' \mapsto \int A_{i,j}(r,z,y',[X_r^{s,\xi,(m-1)}],[X_r^{s,\xi',(m-1)}]) p_m(\mu,s,r,x',y') dy'$ is α-Hölder with modulus bounded by $K(r-s)^{\frac{\eta-\alpha}{2}}$. The previous bound and the mean-value theorem thus yield

$$\begin{aligned} &|\widehat{p}_m(\mu,s,t,x,z) - \widehat{p}_m(\mu',s,t,x,z)| \\ &\leq \frac{K}{t-s} \int_s^t \max_{i,j} |a_{i,j}(r,z,[X_r^{s,\xi,(m-1)}]) - a_{i,j}(r,z,[X_r^{s,\xi',(m-1)}])|\, dr\, g(c(t-s), z-x) \\ &\leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{\alpha-\eta}{2}}}\, g(c(t-s), z-x) \end{aligned}$$

More generally, differentiating $x \mapsto \widehat{p}_m(\mu,s,t,x,z)$, from the mean-value theorem and (5.60), we also obtain

$$\forall \alpha \in [0,1], \quad |\partial_x^n \widehat{p}_m(\mu,s,t,x,z) - \partial_x^n \widehat{p}_m(\mu',s,t,x,z)| \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{n+\alpha-\eta}{2}}}\, g(c(t-s), z-x) \tag{5.61}$$

for $n = 0,1,2$. Now, using the decomposition $p_m(\mu,s,t,x,z) = \widehat{p}_m(\mu,s,t,x,z) + R_m(\mu,s,t,x,z)$ with $R_m(\mu,s,t,x,z) := \sum_{k\geq 1} \widehat{p}_m \otimes \mathcal{H}_m^{(k)}(\mu,s,t,x,z)$ satisfying $|R_m(\mu,s,t,x,z)| \leq K(t-s)^{\eta/2} g(c(t-s), z-x)$,

still in the case $W_2(\mu,\mu') > (t-s)^{\frac{1}{2}}$, we obtain

$$|p_m(\mu,s,t,x,z) - p_m(\mu',s,t,x,z)| \le |\widehat{p}_m(\mu,s,t,x,z) - \widehat{p}_m(\mu',s,t,x,z)| + K(t-s)^{\frac{\eta}{2}} g(c(t-s),z-x)$$
$$\le K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{\alpha-\eta}{2}}} g(c(t-s),z-x).$$

This last estimate concludes the proof of (5.59). We now make use of the following decomposition

$$\mathcal{H}_{m+1}(\mu,s,r,t,y,z) - \mathcal{H}_{m+1}(\mu',s,r,t,y,z) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV}$$

with

$$\mathrm{I} := -\sum_{i=1}^d [b_i(r,y,[X_r^{s,\xi,(m)}]) - b_i(r,y,[X_r^{s,\xi',(m)}])] H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-y \right) \widehat{p}_{m+1}(\mu,s,r,t,y,z),$$

$$\mathrm{II} := -\sum_{i=1}^d b_i(r,y,[X_r^{s,\xi',(m)}]) \Big[H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-y \right) - H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi',(m)}])dv, z-y \right) \Big]$$
$$\times \widehat{p}_{m+1}(\mu,s,r,t,y,z),$$

$$\mathrm{III} := \frac{1}{2} \sum_{i,j=1}^d \Big[(a_{i,j}(r,y,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}]))$$
$$- (a_{i,j}(r,y,[X_r^{s,\xi',(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi',(m)}])) \Big] H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-y \right) \widehat{p}_{m+1}(\mu,s,r,t,y,z),$$

$$\mathrm{IV} := \frac{1}{2} \sum_{i,j=1}^d (a_{i,j}(r,y,[X_r^{s,\xi',(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi',(m)}]))$$
$$\times \Big[H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-y \right) - H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi',(m)}])dv, z-y \right) \Big] \widehat{p}_{m+1}(\mu,s,r,t,y,z),$$

$$\mathrm{V} := -\sum_{i=1}^d b_i(r,y,[X_r^{s,\xi',(m)}]) H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv, z-y \right) \Delta_\mu \widehat{p}_{m+1}(\mu,s,r,t,y,z)$$
$$+ \frac{1}{2} \sum_{i,j=1}^d (a_{i,j}(r,y,[X_r^{s,\xi',(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi',(m)}])) H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi',(m)}])dv, z-y \right)$$
$$\times \Delta_\mu \widehat{p}_{m+1}(\mu,s,r,t,y,z).$$

From similar arguments as those previously used, for all $r \neq s$ and for all $\alpha \in [0,1]$, we get

$$|\mathrm{I}| \le K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y),$$
$$|\mathrm{II}| \le K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y),$$
$$|\mathrm{III}| \le K \left\{ \frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{\alpha}{2}}} \wedge \frac{1}{(t-r)(r-s)^{\frac{\alpha-\eta}{2}}} \right\} W_2^\alpha(\mu,\mu')\, g(c(t-r),z-x),$$
$$|\mathrm{IV}| \le K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y),$$
$$|\mathrm{V}| \le K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y).$$

We only prove the estimates on I and III and omit the remaining technical details. Since each b_i satisfies $(\mathbf{HR}_+)$ we get

$$\begin{aligned}
|\mathrm{I}| &\leq \frac{K}{(t-r)^{\frac{1}{2}}}\Big|\int B_i(r,y,z,[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}])(p_m(\mu,s,r,z')-p_m(\mu',s,r,z')dz'\Big| g(c(t-r),z-y)\\
&\leq \frac{K}{(t-r)^{\frac{1}{2}}}\Big|\int\int B_i(r,y,z',[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}])p_m(\mu,s,r,x',z')\,dz'\,(\mu-\mu')(dx')\Big| g(c(t-r),z-y)\\
&+ \frac{K}{(t-r)^{\frac{1}{2}}}\Big|\int\int B_i(r,y,z',[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}])(p_m(\mu,s,r,x',z')-p_m(\mu',s,r,x',z'))\,dz'\,\mu'(dx')\Big|\\
&\quad\times g(c(t-r),z-y)\\
&\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{1}{2}}(r-s)^{\frac{\alpha-\eta}{2}}}\,g(c(t-r),z-y)
\end{aligned}$$

where we used the estimate (5.61) with $n=0$ and also the fact that, since $z'\mapsto B_i(r,y,z',\mu,\nu)$ is η-Hölder uniformly w.r.t the other variables, the map $x\mapsto\int B_i(r,y,z',[X_r^{s,\xi,(m)}],[X_r^{s,\xi',(m)}])p_m(\mu,s,r,x,z')dz'$ is α-Hölder, for any $\alpha\in[0,1]$ with a modulus bounded by $K(r-s)^{\frac{\eta-\alpha}{2}}$, $K:=K(T,a,b)$ being a positive constant independent of m. From the following identity

$$\begin{aligned}
h(y) &:= a_{i,j}(r,y,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,y,[X_r^{s,\xi',(m)}])\\
&= \int_0^1 \partial_\lambda a_{i,j}(r,y,[X_r^{s,\xi+\lambda(\xi-\xi'),(m))}])\,d\lambda\\
&= \int_0^1 \mathbb{E}\Big[\partial_\mu[\int \widetilde{a}_{i,j}(r,y,y',[X_r^{s,\xi+\lambda(\xi-\xi'),(m)}])p_m(\mu,s,r,y')dy')]_{|\mu=[\xi+\lambda(\xi-\xi')]}(\xi+\lambda(\xi-\xi'))(\xi-\xi')\Big]
\end{aligned}$$

and the η-Hölder regularity of $y\mapsto\widetilde{a}_{i,j}(r,y,z,\mu)$, we deduce

$$|h(y)-h(z)|\leq K\frac{|y-z|^\eta}{(r-s)^{\frac{1}{2}}}\mathbb{E}[|\xi-\xi'|]. \tag{5.62}$$

Taking infimum over all joint distributions of the random variables ξ and ξ' with marginals μ and μ' respectively and plugging the corresponding bound in |III|, we get

$$|\mathrm{III}|\leq K\frac{W_2(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{1}{2}}}\,g(c(t-r),z-y)\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{\alpha}{2}}}\,g(c(t-r),z-y)$$

when $W_2(\mu,\mu')\leq(r-s)^{\frac{1}{2}}$. If $W_2(\mu,\mu')>(r-s)^{\frac{1}{2}}$, we directly get

$$|\mathrm{III}|\leq K\frac{1}{(t-r)^{1-\frac{\eta}{2}}}\,g(c(t-r),z-y)\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{\alpha}{2}}}\,g(c(t-r),z-y).$$

Then, similarly to I, using that each $a_{i,j}$ satisfies $(\mathbf{HR})$, we get

$$|\mathrm{III}|\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)(r-s)^{\frac{\alpha-\eta}{2}}}\,g(c(t-r),z-y).$$

This last bound completes the proof of the third estimate. Gathering all the previous estimates together, we thus obtain

$$\begin{aligned}
&\forall\alpha\in[0,1],\quad |\mathcal{H}_{m+1}(\mu,s,r,t,y,z)-\mathcal{H}_{m+1}(\mu',s,r,t,y,z)|\\
&\qquad\leq K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{\frac{\alpha}{2}}}\wedge\frac{1}{(t-r)(r-s)^{\frac{\alpha-\eta}{2}}}\right\}W_2^\alpha(\mu,\mu')\,g(c(t-r),z-y).
\end{aligned}\tag{5.63}$$

For $r=s$, using a similar decomposition as the one employed above and omitting the remaining technical details, we also obtain

$$\forall\alpha\in[0,\eta),\quad |\mathcal{H}_{m+1}(\mu,s,t,y,z)-\mathcal{H}_{m+1}(\mu',s,t,y,z)|\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-s)^{1+\frac{\alpha-\eta}{2}}}\,g(c(t-s),z-y). \tag{5.64}$$

For a fixed $\mu'\in\mathcal{P}_2(\mathbb{R}^d)$ and $m\geq1$, we introduce the following notations

$$\Delta_\mu\widehat{p}_m(\mu,s,t,x,z):=\widehat{p}_m(\mu,s,t,x,z)-\widehat{p}_m(\mu',s,t,x,z)$$

and proceed similarly with other quantities. Hence, for example, $\Delta_\mu\mathcal{H}_m(\mu,s,t,x,z):=\mathcal{H}_m(\mu,s,t,x,z)-\mathcal{H}_m(\mu',s,t,x,z)$, $\Delta_\mu\partial_v\partial_\mu[\widehat{p}_m(\mu,s,t,x,z)](v)=\partial_v\partial_\mu[\widehat{p}_m(\mu,s,t,x,z)](v)-\partial_v\partial_\mu[\widehat{p}_m(\mu,s,t,x,z)]_{|\mu=\mu'}(v)$

and so on. With the previous notation and our computations, by induction on k, it follows that for any $\alpha \in [0,1]$

$$|\Delta_\mu \widehat{p}_{m+1} \otimes \mathcal{H}^{(k)}_{m+1}(\mu,s,t,x,z)| \leq K^k W_2^\alpha(\mu,\mu')(t-s)^{\frac{(\eta-\alpha)}{2}+k\frac{\eta}{2}} \prod_{i=1}^k B\left(\frac{\eta}{2}, 1+\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right) \times g(c(t-s),z-x),$$

$$|(p_{m+1} \otimes \Delta_\mu \mathcal{H}_{m+1}) \otimes \mathcal{H}^{(k)}_{m+1}(\mu,s,t,x,z)| \leq K^k W_2^\alpha(\mu,\mu')(t-s)^{\frac{(\eta-\alpha)}{2}+k\frac{\eta}{2}} \prod_{i=1}^k B\left(\frac{\eta}{2}, 1+\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right) \times g(c(t-s),z-x).$$

From the representation in infinite series of p_{m+1}, the following relation holds

$$\Delta_\mu p_{m+1}(\mu,s,t,x,z) = \Delta_\mu \widehat{p}_{m+1}(\mu,s,t,x,z) + p_{m+1} \otimes \Delta_\mu \mathcal{H}_{m+1}(\mu,s,t,x,z) + \Delta_\mu p_{m+1} \otimes \mathcal{H}_{m+1}(\mu,s,t,x,z)$$

which in turn by iteration yields

$$\Delta_\mu p_{m+1}(\mu,s,t,x,z) = \sum_{k\geq 0} \{\Delta_\mu \widehat{p}_{m+1} + p_{m+1} \otimes \Delta_\mu \mathcal{H}_{m+1}\} \otimes \mathcal{H}^{(k)}_m(\mu,s,t,x,z). \tag{5.65}$$

Moreover, one may differentiate the infinite series (5.65) with respect to x so that for $n = 0,1,2$

$$\partial_x^n \Delta_\mu p_{m+1}(\mu,s,t,x,z) = \sum_{k\geq 0} \{\partial_x^n \Delta_\mu \widehat{p}_{m+1} + \partial_x^n p_{m+1} \otimes \Delta_\mu \mathcal{H}_{m+1}\} \otimes \mathcal{H}^{(k)}_{m+1}(\mu,s,t,x,z). \tag{5.66}$$

In order to make the previous formula rigorous, we study the iterated kernels that appear in the previous series. From the estimates (5.61) and (5.63) we get

$$|\partial_x \Delta_\mu \widehat{p}_{m+1} \otimes \mathcal{H}^{(k)}_{m+1}(\mu,s,t,x,z)| \leq K^k W_2^\alpha(\mu,\mu')(t-s)^{-\frac{1}{2}+\frac{(\eta-\alpha)}{2}+k\frac{\eta}{2}} \prod_{i=1}^k B\left(\frac{\eta}{2}, \frac{1}{2}+\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right) g(c(t-s),z-x)$$

$$|(\partial_x p_{m+1} \otimes \Delta_\mu \mathcal{H}_{m+1}) \otimes \mathcal{H}^{(k)}_{m+1}(\mu,s,t,x,z)| \leq K^k W_2^\alpha(\mu,\mu')(t-s)^{-\frac{1}{2}+\frac{(\eta-\alpha)}{2}+k\frac{\eta}{2}} \prod_{i=1}^k B\left(\frac{\eta}{2}, \frac{1}{2}+\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right) g(c(t-s),z-x)$$

for any $\alpha \in [0,1]$ and

$$|\partial_x^2 \Delta_\mu \widehat{p}_{m+1} \otimes \mathcal{H}^{(k)}_{m+1}(\mu,s,t,x,z)| \leq K^k W_2^\alpha(\mu,\mu')(t-s)^{-1+\frac{(\eta-\alpha)}{2}+k\frac{\eta}{2}} \prod_{i=1}^k B\left(\frac{\eta}{2}, \frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right) g(c(t-s),z-x)$$

for any $\alpha \in [0,\eta)$. On the one hand, if $r \in [s,(t+s)/2]$, from (5.63), we get

$$\int_{\mathbb{R}^d} |\partial_x^2 p_m(\mu,s,r,x,y)| \Delta_\mu \mathcal{H}_m(\mu,s,r,t,y,z)| dy \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} g(c(t-s),z-x) \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)(r-s)^{1+\frac{\alpha-\eta}{2}}} g(c(t-s),z-x)$$

so that for any $\alpha \in [0,\eta)$

$$\int_s^{\frac{t+s}{2}} \int_{\mathbb{R}^d} |\partial_x^2 p_m(\mu,s,r,x,y)||\Delta_\mu \mathcal{H}_m(\mu,s,r,t,y,z)|\, dy\, dr \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-s)^{1+\frac{\alpha-\eta}{2}}} g(c(t-s),z-x). \tag{5.67}$$

On the other hand, if $r \in [\frac{t+s}{2}, t]$, again from (5.63), we get

$$\int_{\mathbb{R}^d} |\partial_x^2 p_m(\mu,s,r,x,y)||\Delta_\mu \mathcal{H}_m(\mu,s,r,t,y,z)| dy \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha}{2}}} g(c(t-s),z-x) \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(t-s)^{1+\frac{\alpha}{2}}} g(c(t-s),z-x)$$

so that

$$\int_{\frac{t+s}{2}}^{t}\int_{\mathbb{R}^d}|\partial_x^2 p_m(\mu,s,r,x,y)||\Delta_\mu\mathcal{H}_m(\mu,s,r,t,y,z)|\,dy\,dr\le K\frac{W_2^\alpha(\mu,\mu')}{(t-s)^{1+\frac{\alpha-\eta}{2}}}g(c(t-s),z-x). \tag{5.68}$$

Combining (5.67) with (5.68), we finally conclude

$$\forall\alpha\in[0,\eta),\quad |(\partial_x^2 p_m\otimes\Delta_\mu\mathcal{H}_m)(\mu,s,t,x,z)|\le K\frac{W_2^\alpha(\mu,\mu')}{(t-s)^{1+\frac{\alpha-\eta}{2}}}g(c(t-s),z-x)$$

so that, by induction on k, we obtain

$$\begin{aligned}|(\partial_x^2 p_m\otimes\Delta_\mu\mathcal{H}_m)\otimes\mathcal{H}_m^{(k)}(\mu,s,t,x,z)|&\le K^k W_2^\alpha(\mu,\mu')(t-s)^{-1+\frac{(\eta-\alpha)}{2}+k\frac{\eta}{2}}\prod_{i=1}^{k}B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)\\&\quad\times g(c(t-s),z-x)\end{aligned}$$

for any $\alpha\in[0,\eta)$. From the asymptotics of the Beta function, we conclude that the infinite series appearing in the right-hand side of (5.66) is absolutely convergent in the two following cases: $n=0$, 1 for any $\alpha\in[0,1]$ and $n=2$ for any $\alpha\in[0,\eta)$. Moreover, there exist two constants $K:=K(T,a,b)$, $c:=c(\lambda)$ such that for any $\mu,\mu'\in\mathcal{P}_2(\mathbb{R}^d)$

$$|\partial_x^n p_m(\mu,s,t,x,z)-\partial_x^n p_m(\mu',s,t,x,z)|\le K\frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{n+\alpha-\eta}{2}}}\,g(c(t-s),z-x),\ n=0,1,2. \tag{5.69}$$

This completes the proof of (5.16).

In order to obtain (5.17) we proceed as for the previous estimates. To lighten the notations, we introduce the quantities

$$\begin{aligned}u_m^n(s,t)&:=\sup_{(y,\mu,\mu')\in\mathbb{R}^d\times(\mathcal{P}_2(\mathbb{R}^d))^2,\mu\neq\mu'}\int\int(|y''-x'|^\eta\wedge1)\\&\quad\times\frac{|\partial_y^n[\partial_\mu p_m(\mu,s,t,x,y'')](y)-\partial_y^n[\partial_\mu p_m(\mu',s,t,x,y'')](y)|}{W_2^\alpha(\mu,\mu')}\,dy''\,\mu'(dx'),\\v_m^n(s,t)&:=\sup_{(y,\mu,\mu')\in\mathbb{R}^d\times(\mathcal{P}_2(\mathbb{R}^d))^2,\mu\neq\mu'}\int\int\frac{|\partial_y^n[\partial_\mu p_m(\mu,s,t,x',y'')](y)-\partial_y^n[\partial_\mu p_m(\mu',s,t,x',y'')](y)|}{W_2^\alpha(\mu,\mu')}\,dy''\,\mu'(dx')\end{aligned}$$

for a fixed $n=0$, 1, $\alpha\in[0,1]$ for $n=0$ and $\alpha\in[0,\eta)$ if $n=1$. We prove by induction the following key inequalities:

$$u_m^n(s,t)\le C_{m,n}(s,t)(t-s)^{-(\frac{1+n+\alpha}{2}-\eta)}\quad\text{and}\quad v_m^n(s,t)\le C_{m,n}(s,t)(t-s)^{-(\frac{1+n+\alpha-\eta}{2})},$$

with $C_{m,n}(s,t):=\sum_{k=1}^m C^k\prod_{i=1}^k B\left(\frac{\eta}{2},\frac{1-n+\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{(k-1)\frac{\eta}{2}}$. The result being straightforward for $m=1$, we assume that it holds at step m. We first claim

$$\begin{aligned}&\forall\alpha\in[0,\eta),\quad|\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y)-\partial_v^n[\partial_\mu\widehat{p}_{m+1}(\mu',s,r,t,x,z)](y)|\\&\qquad\le KW_2^\alpha(\mu,\mu')\left(\left\{\frac{1}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}}1_{\{r=s\}}+\frac{1}{(r-s)^{\frac{1+n+\alpha-\eta}{2}}}1_{\{r>s\}}\right\}\right.\\&\qquad\quad\left.+\frac{1}{t-r}\int_r^t u_m^n(s,v)dv\right)g(c(t-r),z-x).\end{aligned} \tag{5.70}$$

In order to prove the previous inequality, we make use of the following decomposition:

$$\Delta_\mu\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y)=\mathrm{I}(y)+\mathrm{II}(y)+\mathrm{III}(y)+\mathrm{IV}(y),$$

with

$$\mathrm{I}(y) := \left\{ Df_{z-x}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])dv\right) - Df_{z-x}\left(\int_r^t a(v,z,[X_v^{s,\xi',(m)}])dv\right)\right\}$$
$$\cdot \int_r^t \left\{ \int (\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,v,[X_v^{s,\xi,(m)}]))\partial_x^{1+n} p_m(\mu,s,v,y,y')dy' \right.$$
$$\left. + \int \mu(dx') \int (\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,x',[X_v^{s,\xi,(m)}]))\partial_y^n \partial_\mu [p_m(\mu,s,v,x',y')](y)dy' \right\} dv,$$

$$\mathrm{II}(y) := Df_{z-x}\left(\int_r^t a(v,z,[X_r^{s,\xi',(m)}])dr\right)$$
$$\cdot \int_r^t \left\{ \int (\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,y',[X_v^{s,\xi',(m)}]))\partial_x^{1+n} p_m(\mu,s,v,y,y')\, dy' \right.$$
$$\left. + \int (\widetilde{a}(v,z,y',[X_v^{s,\xi',(m)}]) - \widetilde{a}(v,z,y,[X_v^{s,\xi',(m)}]))(\partial_x^{1+n} p_m(\mu,s,v,y,y') - \partial_x^{1+n} p_m(\mu',s,v,y,y'))dy' \right\} dv,$$

$$\mathrm{III}(y) := Df_{z-x}\left(\int_r^t a(v,z,[X_v^{s,\xi',(m)}])dv\right)$$
$$\cdot \int_r^t \left\{ \int (\mu-\mu')(dx') \int \widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}])\partial_y^n [\partial_\mu p_m(\mu,s,v,x',y')](y)\, dy' \right.$$
$$\left. + \int \mu'(dx') \int (\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,y',[X_v^{s,\xi',(m)}]))\partial_y^n [\partial_\mu p_m(\mu,s,v,x',y')](y)\, dy' \right\} dv,$$

$$\mathrm{IV}(y) := Df_{z-x}\left(\int_r^t a(v,z,[X_v^{s,\xi',(m)}])dv\right) \cdot \int_r^t \int \mu'(dx') \int (\widetilde{a}(v,z,y',[X_v^{s,\xi',(m)}]) - \widetilde{a}(v,z,x',[X_v^{s,\xi',(m)}]))$$
$$\cdot (\partial_y^n [\partial_\mu p_m(\mu,s,v,x',y')](y) - \partial_y^n [\partial_\mu [p_m(\mu',s,v,x',y')](y))dy'dv.$$

From the mean-value theorem and the estimates (5.60), (5.6) and (5.12), there exists a constant $K := K(T,a,b)$ independent of m such that

$$|\mathrm{I}(y)| \le \frac{K}{(t-r)^2} \int_r^t \max_{i,j} |a_{i,j}(v,z,[X_v^{s,\xi,(m)}]) - a_{i,j}(v,z,[X_v^{s,\xi',(m)}])|\, dv$$
$$\times \left(\int_r^t \left\{ (v-s)^{-\frac{(1+n-\eta)}{2}} + (v-s)^{-\frac{(1+n-\eta)}{2} - \frac{\eta}{2}} \right\} dv \right) g(c(t-r), z-x)$$
$$\le \frac{K}{(t-r)^2} \int_r^t \frac{W_2^\alpha(\mu,\mu')}{(v-s)^{\frac{\alpha-\eta}{2}}}\, dv \times [(t-s)^{1+\frac{1+n-\eta}{2}} 1_{\{r=s\}} + (t-r)(r-s)^{-\frac{1+n-\eta}{2}} 1_{\{r>s\}}] g(c(t-r), z-x)$$
$$\le K W_2^\alpha(\mu,\mu') \left[\frac{1}{(t-s)^{\frac{1+n+\alpha}{2} - \eta}} 1_{\{r=s\}} + \frac{1}{(r-s)^{\frac{1+n+\alpha}{2} - \eta}} 1_{\{r>s\}} \right] g(c(t-r), z-x)$$

for any $\alpha \in [0,1]$. From $(\mathbf{CS}_+)_2$, similarly to (5.60) with the map $\widetilde{A}_{i,j}$ instead of $A_{i,j}$, one gets

$$\forall \alpha \in [0,1],\ |\widetilde{a}(v,z,y'',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,y'',[X_v^{s,\xi',(m)}])| \le K\, W_2^\alpha(\mu,\mu')(v-s)^{\frac{\eta-\alpha}{2}} \tag{5.71}$$

and by (5.6) we obtain

$$\left| \int (\widetilde{a}(v,z,y'',[X_v^{s,\xi,(m)}]) - \widetilde{a}(v,z,y'',[X_v^{s,\xi',(m)}]))\partial_x^{1+n} p_m(\mu,s,v,y,y'')\, dy'' \right| \le K \frac{W_2^\alpha(\mu,\mu')}{(v-s)^{\frac{1+n+\alpha-\eta}{2}}}$$

and, from (5.69), the η-Hölder regularity of $y \mapsto \widetilde{a}(v,z,y,\mu)$ and the space-time inequality (1.4), one has

$$\left| \int (\widetilde{a}(v,z,y'',[X_v^{s,\xi',(m)}]) - \widetilde{a}(v,z,y,[X_v^{s,\xi',(m)}]))(\partial_x^{1+n} p_m(\mu,s,v,y,y'') - \partial_x^{1+n} p_m(\mu',s,v,y,y''))dy'' \right|$$
$$\le K \frac{W_2^\alpha(\mu,\mu')}{(v-s)^{\frac{1+n+\alpha}{2} - \eta}}$$

for any $\alpha \in [0,1]$ if $n = 0$ and for any $\alpha \in [0,\eta)$ if $n = 1$. Combining the two previous estimates, we finally obtain

$$|\mathrm{II}(y)| \le K W_2^\alpha(\mu,\mu') \left[\frac{1}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}} 1_{\{r=s\}} + \frac{1}{(r-s)^{\frac{1+n+\alpha}{2} - \eta}} 1_{\{r>s\}} \right] g(c(t-r), z-x)$$

with $\alpha \in [0,1]$ if $n=0$ and $\alpha \in [0,\eta)$ if $n=1$. From the representation in infinite series (5.40), we derive that $x \mapsto \partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)(y)$ is α-Hölder continuous, with $\alpha \in [0,1]$ if $n=0$ and $\alpha \in [0,\eta)$ if $n=1$. More precisely, for any $x,x' \in \mathbb{R}^d$ and for any $m \geq 1$, one has

$$\begin{aligned}&\left|\partial_v^n[\partial_\mu p_m(\mu,s,t,x,z)](y) - \partial_v^n[\partial_\mu p_m(\mu,s,t,x',z)](y)\right| \\ &\qquad \leq K\frac{|x-x'|^\alpha}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}}\left\{g(c(t-s),z-x)+g(c(t-s),z-x')\right\}. \end{aligned} \tag{5.72}$$

The previous estimate in turn implies that $x \mapsto \int \widetilde{a}(v,z,y'',[X_v^{s,\xi,(m)}])\partial_y^n[\partial_\mu p_m(\mu,s,v,x,y'')](y)dy'' = \int(\widetilde{a}(v,z,y'',[X_v^{s,\xi,(m)}])-\widetilde{a}(v,z,x_0,[X_v^{s,\xi,(m)}]))\partial_y^n[\partial_\mu p_m(\mu,s,v,x,y'')](y)dy''$, $x_0 \in \mathbb{R}^d$, is α-Hölder, with a modulus bounded by $K(v-s)^{-\frac{(1+n+\alpha)}{2}+\eta}$, K being a positive constant independent of m, where $\alpha \in [0,1]$ of $n=0$ and $\alpha \in [0,\eta)$ if $n=1$. We thus deduce

$$\Big|\int\int \widetilde{a}(v,y,y',[X_v^{s,\xi,(m)}])\,\partial_y^n[\partial_\mu p_m(\mu,s,v,x',y')](y)\,dy'\,(\mu-\mu')(dx')\Big| \leq K\frac{W_2^\alpha(\mu,\mu')}{(v-s)^{\frac{1+n+\alpha}{2}-\eta}}.$$

From (5.90) and (5.12), we also obtain

$$|\int\int(\widetilde{a}(v,z,y',[X_v^{s,\xi,(m)}])-\widetilde{a}(v,z,y',[X_v^{s,\xi',(m)}]))\partial_y^n[\partial_\mu p_m(\mu,s,v,x',y')](y)\,dy'\,\mu'(dx')| \leq K\frac{W_2^\alpha(\mu,\mu')}{(v-s)^{\frac{1+n+\alpha}{2}-\eta}}$$

for any $\alpha \in [0,1]$. Consequently, combining the two previous estimates, we conclude

$$|\mathrm{III}(y)| \leq KW_2^\alpha(\mu,\mu')\left\{\frac{1}{(t-s)^{\frac{1+n+\alpha}{2}-\eta}}1_{\{r=s\}}+\frac{1}{(r-s)^{\frac{1+n+\alpha}{2}-\eta}}1_{\{r>s\}}\right\}g(c(t-r),z-x)$$

for any $\alpha \in [0,1]$ if $n=0$ and for any $\alpha \in [0,\eta)$ if $n=1$. Finally, one has

$$|\mathrm{IV}(y)| \leq \frac{K}{t-r}\int_r^t\int\int(|y''-x'|^\eta \wedge 1)|\Delta_\mu\partial_v^n\partial_\mu[p_m(\mu,s,v,x',y'')](v)|\,dy''\,dv\;\mu'(dx')\,g(c(t-r),z-x).$$

Gathering the estimates on $\mathrm{I}(y)$, $\mathrm{II}(y)$, $\mathrm{III}(y)$ and $\mathrm{IV}(y)$, we obtain

$$\begin{aligned}|\Delta_\mu\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y)| &\leq K\left\{W_2^\alpha(\mu,\mu')\left\{\frac{1}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}}1_{\{r=s\}}+\frac{1}{(r-s)^{\frac{1+n+\alpha-\eta}{2}}}1_{\{r>s\}}\right\}\right. \\ &\quad\left.+\frac{1}{t-r}\int_r^t\int\int(|y''-x'|^\eta\wedge 1)|\Delta_\mu\partial_y^n[\partial_\mu p_m(\mu,s,v,x',y'')](y)|\,dy''\,\mu'(dx')\,dv\right\} \\ &\quad\times g(c(t-r),z-x)\end{aligned} \tag{5.73}$$

for any $\alpha \in [0,1]$ if $n=0$ and for any $\alpha \in [0,\eta)$ if $n=1$. This completes the proof of (5.70). As a consequence, from the induction hypothesis, we directly get

$$\begin{aligned}&|\Delta_\mu\partial_v^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,t,x,z)](v)| \\ &\quad\leq K\left\{\frac{1}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}}+\frac{1}{t-s}\int_s^t\frac{C_{m,n}(s,r)}{(r-s)^{\frac{1+n+\alpha}{2}-\eta}}dr\right\}W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x) \\ &\quad\leq\frac{K}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}}\left\{B\left(\frac{\eta}{2},\frac{1-n+\alpha-\eta}{2}\right)+\sum_{k=1}^m C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\} \\ &\qquad\times W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x)\end{aligned}$$

which in turn yields

$$\begin{aligned}&\sum_{r\geq 0}|\Delta_\mu\partial_v^n\partial_\mu\widehat{p}_{m+1}\otimes\mathcal{H}_{m+1}^{(r)}(\mu,s,t,x,z)(v)| \\ &\qquad\leq K\left\{B\left(\frac{\eta}{2},\frac{1-n+\alpha-\eta}{2}\right)+\sum_{k=1}^m C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\} \\ &\qquad\times\frac{W_2^\alpha(\mu,\mu')}{(t-s)^{\frac{1+n+\alpha-\eta}{2}}}g(c(t-s),z-x).\end{aligned} \tag{5.74}$$

From (5.38), we easily obtain the following decomposition

$$\Delta_\mu\partial_v[\partial_\mu\mathcal{H}_{m+1}(\mu,s,r,t,y,z)](v) = \mathrm{A}+\mathrm{B}+\mathrm{C}+\mathrm{D}+\mathrm{E}$$

with

$$
\begin{aligned}
\mathrm{A} &:= \mathrm{A}_1 + \mathrm{A}_2 + \mathrm{A}_3,\\
\mathrm{A}_1 &:= -\sum_{i=1}^{d} \Delta_\mu \partial_v [\partial_\mu b_i(r, y, [X_r^{s,\xi,(m)}])](v)\, H_1^i \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{A}_2 &:= -\sum_{i=1}^{d} \partial_v [\partial_\mu b_i(r, y, [X_r^{s,\xi',(m)}])](v)\, \Delta_\mu H_1^i \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{A}_3 &:= -\sum_{i=1}^{d} \partial_v [\partial_\mu b_i(r, y, [X_r^{s,\xi',(m)}])](v)\, H_1^i \left(\int_r^t a(v', z, [X_{v'}^{s,\xi',(m)}]) dv', z - y \right) \Delta_\mu \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{B} &:= \mathrm{B}_1 + \mathrm{B}_2 + \mathrm{B}_3,\\
\mathrm{B}_1 &:= \frac{1}{2} \sum_{i,j=1}^{d} \Delta_\mu \partial_v [\partial_\mu [a_{i,j}(r, y, [X_r^{s,\xi,(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi,(m)}])]](v)\\
&\qquad \times H_2^{i,j} \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{B}_2 &:= \frac{1}{2} \sum_{i,j=1}^{d} \partial_v [\partial_\mu [a_{i,j}(r, y, [X_r^{s,\xi',(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi',(m)}])]](v)\\
&\qquad \times \Delta_\mu H_2^{i,j} \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{B}_3 &:= \frac{1}{2} \sum_{i,j=1}^{d} \partial_v [\partial_\mu [a_{i,j}(r, y, [X_r^{s,\xi',(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi',(m)}])]](v)\\
&\qquad \times H_2^{i,j} \left(\int_r^t a(v', z, [X_{v'}^{s,\xi',(m)}]) dv', z - y \right) \Delta_\mu \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{C} &:= \mathrm{C}_1 + \mathrm{C}_2 + \mathrm{C}_3,\\
\mathrm{C}_1 &:= -\sum_{i=1}^{d} \Delta_\mu b_i(r, y, [X_r^{s,\xi,(m)}])\, \partial_v \left[\partial_\mu H_1^i \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \right](v)\, \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{C}_2 &:= -\sum_{i=1}^{d} b_i(r, y, [X_r^{s,\xi',(m)}])\, \Delta_\mu \partial_v \left[\partial_\mu H_1^i \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \right](v)\, \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{C}_3 &:= -\sum_{i=1}^{d} b_i(r, y, [X_r^{s,\xi',(m)}])\, \partial_v \left[\partial_\mu H_1^i \left(\int_r^t a(v', z, [X_{v'}^{s,\xi',(m)}]) dv', z - y \right) \right](v)\, \Delta_\mu \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{D} &:= \mathrm{D}_1 + \mathrm{D}_2 + \mathrm{D}_3,\\
\mathrm{D}_1 &:= \frac{1}{2} \sum_{i,j=1}^{d} \Delta_\mu \Big(a_{i,j}(r, y, [X_r^{s,\xi,(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi,(m)}]) \Big)\\
&\qquad \times \partial_v \left[\partial_\mu H_2^{i,j} \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \right](v)\, \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{D}_2 &:= \frac{1}{2} \sum_{i,j=1}^{d} (a_{i,j}(r, y, [X_r^{s,\xi',(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi',(m)}]))\\
&\qquad \times \Delta_\mu \partial_v \left[\partial_\mu H_2^{i,j} \left(\int_r^t a(v', z, [X_{v'}^{s,\xi,(m)}]) dv', z - y \right) \right](v)\, \widehat{p}_{m+1}(\mu, s, r, t, y, z),\\
\mathrm{D}_3 &:= \frac{1}{2} \sum_{i,j=1}^{d} (a_{i,j}(r, y, [X_r^{s,\xi',(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi',(m)}]))\\
&\qquad \times \partial_v \left[\partial_\mu H_2^{i,j} \left(\int_r^t a(v', z, [X_{v'}^{s,\xi',(m)}]) dv', z - y \right) \right](v)\, \Delta_\mu \widehat{p}_{m+1}(\mu, s, r, t, y, z),
\end{aligned}
$$

and

$$\mathrm{E} := \mathrm{E}_1 + \mathrm{E}_2 + \mathrm{E}_3$$

$$\mathrm{E}_1 := -\sum_{i=1}^{d} \Delta_\mu \Big[b_i(r,y,[X_r^{s,\xi,(m)}]) H_1^i \left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y \right) \Big] \partial_v[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,y,z)](v),$$

$$\mathrm{E}_2 := \frac{1}{2} \sum_{i,j=1}^{d} \Delta_\mu \Big[\Big(a_{i,j}(r,y,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}]) \Big) H_2^{i,j} \left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y \right) \Big]$$
$$\times \partial_v[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,y,z)](v),$$

$$\mathrm{E}_3 := \Bigg\{ -\sum_{i=1}^{d} b_i(r,y,[X_r^{s,\xi,(m)}]) H_1^i \left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y \right)$$
$$+ \frac{1}{2} \sum_{i,j=1}^{d} \Big(a_{i,j}(r,y,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}]) \Big) H_2^{i,j} \left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y \right) \Bigg\}$$
$$\times \Delta_\mu \partial_v[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,t,y,z)](v).$$

• **Estimates on** A**:**

In order to deal with A_1, we use the following decomposition

$$\Delta_\mu \partial_v[\partial_\mu b_i(r,y,[X_r^{s,\xi,(m)}])](v) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}$$

with

$$\mathrm{I} := \int (\widetilde{b}_i(r,y,z,[X_r^{s,\xi,(m)}]) - \widetilde{b}_i(r,y,z,[X_r^{s,\xi',(m)}])) \, \partial_x^2 p_m(\mu,s,r,v,z) \, dz,$$

$$\mathrm{II} := \int (\widetilde{b}_i(r,y,z,[X_r^{s,\xi',(m)}]) - \widetilde{b}_i(r,y,v,[X_r^{s,\xi',(m)}])) \, (\partial_x^2 p_m(\mu,s,r,v,z) - \partial_x^2 p_m(\mu',s,r,v,z)) \, dz,$$

$$\mathrm{III} := \int \int \widetilde{b}_i(r,y,z,[X_r^{s,\xi,(m)}]) \, \partial_v[\partial_\mu p_m(\mu,s,r,x',z)](v) \, dz \, (\mu-\mu')(dx'),$$

$$\mathrm{IV} := \int \int (\widetilde{b}_i(r,y,z,[X_r^{s,\xi,(m)}]) - \widetilde{b}_i(r,y,z,[X^{s,\xi',(m)}])) \, \partial_v[\partial_\mu p_m(\mu,s,r,x',z)](v) \, dz \, \mu'(dx'),$$

$$\mathrm{V} := \int \int (\widetilde{b}_i(r,y,z,[X_r^{s,\xi',(m)}]) - \widetilde{b}_i(r,y,x',[X_r^{s,\xi',(m)}]))$$
$$.(\partial_v[\partial_\mu p_m(\mu,s,r,x',z)](v) - \partial_v[\partial_\mu p_m(\mu',s,r,x',z)](v)) \, dz \, \mu'(dx').$$

We now need to quantify the contribution of each term appearing above. From $(\mathbf{CS}_+)_2$ and (5.69) (with $n=0$), similarly to (5.60) with $\widetilde{B}_i$ instead of $A_{i,j}$, we directly get $|\widetilde{b}_i(r,y,z,[X_r^{s,\xi,(m)}]) - \widetilde{b}_i(r,y,z,[X_r^{s,\xi',(m)}])| \leq K W_2^\beta(\mu,\mu')(r-s)^{-(\beta-\eta)/2}$, so that

$$|\mathrm{I}| \leq K \frac{W_2^\beta(\mu,\mu')}{(r-s)^{1+\frac{\beta-\eta}{2}}}.$$

From (5.69) and the η-Hölder regularity of $z \mapsto \widetilde{b}_i(r,x,z,\mu)$, we also obtain

$$|\mathrm{II}| \leq K \frac{W_2^\beta(\mu,\mu')}{(r-s)^{1+\frac{\beta}{2}-\eta}}.$$

Using the fact that $x \mapsto \int \widetilde{b}_i(r,y,z,[X_r^{s,\xi,(m)}]) \partial_v[\partial_\mu p_m(\mu,s,r,x,z)](v) \, dz$ is β-Hölder, $\beta \in [0,\eta)$, with modulus bounded by $K(r-s)^{-(1+\frac{\beta}{2}-\eta)}$, we get

$$|\mathrm{III}| \leq K \frac{W_2^\beta(\mu,\mu')}{(r-s)^{1+\frac{\beta}{2}-\eta}}.$$

Similarly to I, from (5.69) (with $n=0$) and (5.12) (with $n=1$), one has

$$|\mathrm{IV}| \leq K \frac{W_2^\beta(\mu,\mu')}{(r-s)^{1+\frac{\beta}{2}-\eta}}.$$

Finally, for the last term, one has

$$|\mathrm{V}| \leq K \int \int (|z-x'|^\eta \wedge 1) |\partial_v[\partial_\mu p_m(\mu,s,r,x',z)](v) - \partial_v[\partial_\mu p_m(\mu',s,r,x',z)](v)| \, dz \, \mu'(dx').$$

Gathering the previous estimates and using the induction hypothesis, we finally obtain

$$|\Delta_\mu \partial_v[\partial_\mu b_i(r,y,[X_r^{s,\xi,(m)}])](v)| \le K\Big(\frac{1}{(r-s)^{1+\frac{\beta-\eta}{2}}} + u_m^n(s,r)\Big)\, W_2^\beta(\mu,\mu')$$

so that

$$|\mathrm{A}_1| \le \frac{K}{(t-r)^{\frac12}(r-s)^{1+\frac{\beta-\eta}{2}}}(1+C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}})\, W_2^\beta(\mu,\mu')\, g(c(t-r),z-y).$$

For A_2, from (5.29) and (5.12) one gets (5.30) so that

$$|\partial_v[\partial_\mu b_i(r,y,[X_r^{s,\xi',(m)}])](v)| \le \frac{K}{(r-s)^{1-\frac{\eta}{2}}} \tag{5.75}$$

and, from the mean-value theorem and (5.60), for any $\alpha\in[0,1]$

$$\begin{aligned}
&\Big|\Delta_\mu H_1^i\Big(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y\Big)\widehat{p}_{m+1}(\mu,s,r,t,y,z)\Big| \\
&\qquad \le \frac{K}{(t-r)^{\frac32}}\int_r^t \max_{i,j}|a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}]) - a_{i,j}(v',z,[X_{v'}^{s,\xi',(m)}])|dv'\, g(c(t-r),z-y) \\
&\qquad \le K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac12}(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y)
\end{aligned}$$

so that

$$\forall \alpha\in[0,1],\quad |\mathrm{A}_2| \le K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac12}(r-s)^{1+\frac{\alpha}{2}-\eta}} g(c(t-r),z-y).$$

For A_3, from (5.60) and the mean value theorem, one similarly gets

$$\Big|\Delta_\mu \widehat{p}_{m+1}(\mu,s,r,t,y,z)\Big| \le K\frac{W_2^\alpha(\mu,\mu')}{(r-s)^{\frac{\alpha-\eta}{2}}}\, g(c(t-r),z-y) \tag{5.76}$$

which in turn, with (5.75), directly imply

$$\forall \alpha\in[0,1],\quad |\mathrm{A}_3| \le K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac12}(r-s)^{1+\frac{\alpha}{2}-\eta}}\, g(c(t-r),z-y).$$

Combining the previous estimates, we finally obtain

$$|\mathrm{A}| \le \frac{K}{(t-r)^{\frac12}(r-s)^{1+\frac{\alpha-\eta}{2}}}(1+C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}})\, W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).$$

• **Estimates on** B**:**

For B_1, we employ a similar decomposition as for $\Delta_\mu\partial_v[\partial_\mu b_i(r,y,[X_r^{s,\xi,(m)}])](v)$, namely

$$\Delta_\mu\partial_v[\partial_\mu a_{i,j}(r,y,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}])](v) = \mathrm{I}_{i,j}+\mathrm{II}_{i,j}+\mathrm{III}_{i,j}+\mathrm{IV}_{i,j}+\mathrm{V}_{i,j}$$

with

$$\begin{aligned}
\mathrm{I}_{i,j} &:= \int \Big[\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi,(m)}]) \\
&\quad - (\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi',(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi',(m)}]))\Big].\partial_x^2 p_m(\mu,s,r,v,z')\,dz', \\
\mathrm{II}_{i,j} &:= \int \Big[\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi',(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi',(m)}]) \\
&\quad - (\widetilde{a}_{i,j}(r,y,v,[X_r^{s,\xi',(m)}]) - \widetilde{a}_{i,j}(r,z,v,[X_r^{s,\xi',(m)}]))\Big].(\partial_x^2 p_m(\mu,s,r,v,z') - \partial_x^2 p_m(\mu',s,r,v,z'))\,dz', \\
\mathrm{III}_{i,j} &:= \int\int \Big[\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi,(m)}])\Big]\partial_v[\partial_\mu p_m(\mu,s,r,x',z')](v)\,dz'\,(\mu-\mu')(dx'), \\
\mathrm{IV}_{i,j} &:= \int\int \Big[\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi,(m)}]) \\
&\quad - (\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi',(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi',(m)}]))\Big].\partial_v[\partial_\mu p_m(\mu,s,r,x',z')](v)\,dz'\,\mu'(dx'), \\
\mathrm{V}_{i,j} &:= \int\int \Big[\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi',(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi',(m)}]) - (\widetilde{a}_{i,j}(r,y,x',[X_r^{s,\xi',(m)}]) - \widetilde{a}_{i,j}(r,z,x',[X_r^{s,\xi',(m)}]))\Big] \\
&\quad .(\partial_v[\partial_\mu p_m(\mu,s,r,x',z')](v) - \partial_v[\partial_\mu p_m(\mu',s,r,x',z')](v))\,dz'\,\mu'(dx').
\end{aligned}$$

As previously done, we quantify the contribution of each term in the above decomposition. First, using condition $(\mathbf{CS}_+)_2$ for the map $\mu \mapsto \widetilde{a}_{i,j}(r,x,y,\mu)$, similarly to (5.60) with the map $\widetilde{A}_{i,j}$ instead of $A_{i,j}$, we directly get

$$\forall \alpha \in [0,1], \quad |\mathrm{I}_{i,j}| \leq K \frac{W_2^\alpha(\mu,\mu')}{(r-s)^{1+\frac{\alpha-\eta}{2}}}.$$

Otherwise, using the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(r,x,z,\mu)$, we get

$$|\mathrm{I}_{i,j}| \leq K \frac{|y-z|^\eta}{r-s}.$$

Hence, combining both estimates with the space-time inequality (1.4), we deduce

$$\begin{aligned} \forall \alpha \in [0,1], \quad & \left|\mathrm{I}_{i,j} \times H_2^{i,j}\left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y\right)\widehat{p}_{m+1}(\mu,s,r,t,y,z)\right| \\ & \leq K\left\{\frac{W_2^\alpha(\mu,\mu')}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)}\right\} g(c(t-r),z-y). \end{aligned} \tag{5.77}$$

Now, if $W_2(\mu,\mu') \geq (t-r)^{1/2}$, then, from the previous bound, we directly get

$$\begin{aligned} \forall \alpha \in [0,1], \quad & |\mathrm{I}_{i,j} \times H_2^{i,j}(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y)\,\widehat{p}_{m+1}(\mu,s,r,t,y,z)| \\ & \leq K \frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)} g(c(t-r),z-y). \end{aligned}$$

Otherwise, if $W_2(\mu,\mu') < (t-r)^{1/2}$, then from (5.77) with $\alpha = \eta$, for any $\beta \in [0,\eta)$,

$$\begin{aligned} & |\mathrm{I}_{i,j} \times H_2^{i,j}(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y)\,\widehat{p}_{m+1}(\mu,r,t,y,z)| \\ & \leq K \frac{W_2^\eta(\mu,\mu')}{(t-r)(r-s)}\, g(c(t-r),z-y) \\ & \leq K \frac{W_2^\beta(\mu,\mu')}{(t-r)^{1+\frac{\beta-\eta}{2}}(r-s)} g(c(t-r),z-y). \end{aligned}$$

Gathering the three previous estimates, we finally obtain

$$\begin{aligned} \forall \alpha \in [0,\eta), \quad & |\mathrm{I}_{i,j} \times H_2^{i,j}(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y)\,\widehat{p}_{m+1}(\mu,r,t,y,z)| \\ & \leq K\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\} W_2^\alpha(\mu,\mu') g(c(t-r),z-y). \end{aligned}$$

Again, from (5.69) and the η-Hölder regularity of $y \mapsto \widetilde{a}_{i,j}(r,y,z,\mu)$, we obtain

$$\forall \alpha \in [0,\eta), \quad |\mathrm{II}_{i,j}| \leq K \frac{|y-z|^\eta \wedge 1}{(r-s)^{1+\frac{\alpha-\eta}{2}}} W_2^\alpha(\mu,\mu').$$

From (5.72), the map $x' \mapsto \int(\widetilde{a}_{i,j}(r,y,z',[X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r,z,z',[X_r^{s,\xi,(m)}]))\partial_v[\partial_\mu p_m(\mu,s,r,x',z')](v)\,dz'$ is α-Hölder with a modulus bounded by $K(|y-z|^\eta \wedge 1)(r-s)^{-1-\frac{(\alpha-\eta)}{2}}$ so that

$$\forall \alpha \in [0,\eta), \quad |\mathrm{III}_{i,j}| \leq K \frac{|y-z|^\eta \wedge 1}{(r-s)^{1+\frac{\alpha-\eta}{2}}} W_2^\alpha(\mu,\mu').$$

Using similar arguments as those employed for $\mathrm{I}_{i,j}$, we get

$$\forall \alpha \in [0,1], \quad |\mathrm{IV}_{i,j}| \leq C\left\{\frac{W_2^\alpha(\mu,\mu')}{(r-s)^{1+\frac{\alpha}{2}-\eta}} \wedge \frac{|y-z|^\eta}{(r-s)^{1-\frac{\eta}{2}}}\right\}$$

so that, considering the two cases $W_2(\mu,\mu') \geq (t-r)^{\frac{1}{2}}$ and $W_2(\mu,\mu') \leq (t-r)^{\frac{1}{2}}$ as previously done,

$$\begin{aligned} \forall \alpha \in [0,\eta), \quad & |\mathrm{IV}_{i,j} \times H_2^{i,j}\left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y\right)\widehat{p}_{m+1}(\mu,r,t,y,z)| \\ & \leq K\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\} W_2^\alpha(\mu,\mu') g(c(t-r),z-y). \end{aligned}$$

For the last term, using the fact that $\mu \mapsto a(r,x,\mu)$ is in $(\mathbf{CS}_+)$, more precisely, the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(t,x,z,\mu)$ on the one hand or the η-Hölder regularity of $z \mapsto \widetilde{a}_{i,j}(t,x,z,\mu)$ on the other hand, as well as the induction hypothesis, we get

$$|\mathrm{V}_{i,j}| \le K\left\{u_m^n(s,r) \wedge |y-z|^\eta v_m^n(s,r)\right\} W_2^\alpha(\mu,\mu').$$

Gathering the previous estimates and using the space-time inequality (1.4), we finally obtain

$$\begin{aligned}|\mathrm{B}_1| &\le K\left\{\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\} + \left\{\frac{u_m^n(s,r)}{t-r} \wedge \frac{v_m^n(s,r)}{(t-r)^{1-\frac{\eta}{2}}}\right\}\right\} W_2^\alpha(\mu,\mu') \\ &\quad \times g(c(t-r),z-y)\end{aligned}$$

for all $\alpha \in [0,\eta)$. For B_2, from (5.46) (bounding $(r-s)^{\frac{1+n-\eta}{2}} u_m^n(s,r)$ and $(r-s)^{\frac{1+n}{2}} v_m^n(s,r)$ by K), one gets

$$|\partial_v[\partial_\mu[a_{i,j}(r,y,[X_r^{s,\xi',(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi',(m)}])]](v)| \le K\left\{\frac{|z-y|^\eta}{r-s} \wedge \frac{1}{(r-s)^{1-\frac{\eta}{2}}}\right\} \tag{5.78}$$

and, by the mean-value theorem and (5.60)

$$\begin{aligned}&|\Delta_\mu H_2^{i,j}\left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv', z-y\right)\widehat{p}_{m+1}(\mu,r,t,y,z)| \\ &\le \frac{K}{(t-r)^2}\int_r^t \max_{i,j}|a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}]) - a_{i,j}(v',z,[X_{v'}^{s,\xi',(m)}])|dv' g(c(t-r),z-y) \\ &\le K\frac{W_2^\alpha(\mu,\mu')}{(t-r)(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y)\end{aligned} \tag{5.79}$$

so that

$$\begin{aligned}|\mathrm{B}_2| &\le K\left\{\frac{1}{(r-s)^{1-\frac{\eta}{2}}} \wedge \frac{|z-y|^\eta}{(r-s)}\right\}\frac{W_2^\alpha(\mu,\mu')}{(t-r)(r-s)^{\frac{\alpha-\eta}{2}}} g(c(t-r),z-y) \\ &\le K\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha}{2}-\eta}} \wedge \frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\right\} W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).\end{aligned}$$

For B_3, from (5.76) and (5.78), we get

$$\begin{aligned}|\mathrm{B}_3| &\le K\left\{\frac{1}{(t-r)(r-s)^{1-\frac{\eta}{2}}} \wedge \frac{|z-y|^\eta}{(t-r)(r-s)}\right\}\frac{W_2^\alpha(\mu,\mu')}{(r-s)^{\frac{\alpha-\eta}{2}}}\, g(c(t-r),z-y) \\ &\le K\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha}{2}-\eta}} \wedge \frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\right\} W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).\end{aligned}$$

Gathering the previous estimates on B_1, B_2, B_3 and using the induction hypothesis, we finally deduce

$$\begin{aligned}|\mathrm{B}| &\le K\left\{\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\} + \left\{\frac{u_m^n(s,r)}{t-r} \wedge \frac{v_m^n(s,r)}{(t-r)^{1-\frac{\eta}{2}}}\right\}\right\} W_2^\alpha(\mu,\mu') \\ &\quad \times g(c(t-r),z-y) \\ &\le K\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\}(1+C_{m,n}(s,r)(r-s)^{\frac{\eta}{2}}) \\ &\quad \times W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).\end{aligned}$$

• **Estimates on C:**

For C_1, from $(\mathbf{HR}_+)$ and (5.69) (with $n=0$), similarly to (5.60) with the map b_i instead of $a_{i,j}$, one has

$$\forall \alpha \in [0,1], \quad |\Delta_\mu b_i(r,y,[X_r^{s,\xi,(m)}])| \le K\frac{W_2^\alpha(\mu,\mu')}{(r-s)^{\frac{\alpha-\eta}{2}}} \tag{5.80}$$

and, from (5.29) (with $n=1$) and (5.12) one gets (5.30) (with $n=1$), so that

$$\begin{aligned}
&\Big|\partial_v\Big[\partial_\mu H_1^i\left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv',z-y\right)\Big](v)\widehat{p}_{m+1}(\mu,r,t,y,z)\Big| \\
&\leq \frac{K}{(t-r)^{\frac{3}{2}}}\int_r^t \max_{i,j}|\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}])](v)|\,dv'\,g(c(t-r),z-y) \qquad (5.81)\\
&\leq \frac{K}{(t-r)^{\frac{1}{2}}(r-s)^{1-\frac{\eta}{2}}}\,g(c(t-r),z-y).
\end{aligned}$$

Combining both estimates, we obtain

$$|\mathrm{C}_1| \leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{\alpha}{2}-\eta}}\,g(c(t-r),z-y).$$

From similar computations as those employed for the term $\Delta_\mu\partial_v[\partial_\mu b_i(r,y,[X_r^{s,\xi,(m)}])](v)$, in a completely analogous manner, we get

$$|\Delta_\mu\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}])](v)| \leq K\Big(\frac{1}{(v'-s)^{1+\frac{\alpha-\eta}{2}}}+u_m^n(s,v')\Big)\,W_2^\alpha(\mu,\mu') \qquad (5.82)$$

so that, from the mean-value theorem and the induction hypothesis

$$\begin{aligned}
|\mathrm{C}_2| &\leq \frac{K}{(t-r)^{\frac{3}{2}}}\int_r^t \max_{i,j}|\Delta_\mu\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m-1)}])](v)|dv'\,g(c(t-r),z-y)\\
&\leq \frac{K}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\left(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv\right)\,W_2^\alpha(\mu,\mu')\,g(c(t-r),z-y).
\end{aligned}$$

For C_3, from (5.76) and then (5.81), we obtain

$$\begin{aligned}
|\mathrm{C}_3| &\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{3}{2}}(r-s)^{\frac{\alpha-\eta}{2}}}\int_r^t \max_{i,j}|\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}])](v)|\,dv'\,g(c(t-r),z-y)\\
&\leq K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{\alpha}{2}-\eta}}\,g(c(t-r),z-y).
\end{aligned}$$

Gathering the previous estimates on C_1, C_2 and C_3, we get

$$|\mathrm{C}| \leq \frac{K}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\left(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv\right)\,W_2^\alpha(\mu,\mu')\,g(c(t-r),z-y).$$

• **Estimates on** D**:**

In order to deal with D_1, we first remark that from (5.62) and the computations shortly after, distinguishing the two cases $W_2(\mu,\mu')\geq(r-s)^{\frac{1}{2}}$ and $W_2(\mu,\mu')\leq(r-s)^{\frac{1}{2}}$, we get

$$\forall\alpha\in[0,1],\quad |\Delta_\mu\Big(a_{i,j}(r,y,[X_r^{s,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s,\xi,(m)}])\Big)| \leq K\frac{|z-y|^\eta}{(r-s)^{\frac{\alpha}{2}}}W_2^\alpha(\mu,\mu').$$

From (5.60), we also get

$$\forall\alpha\in[0,1],\quad |\Delta_\mu a_{i,j}(r,y,[X_r^{s,\xi,(m)}])|+|\Delta_\mu a_{i,j}(r,z,[X_r^{s,\xi,(m)}])| \leq K\frac{W_2^\alpha(\mu,\mu')}{(r-s)^{\frac{\alpha-\eta}{2}}}.$$

Gathering the two previous bounds, one obtains

$$\forall\alpha\in[0,1],\quad |\Delta_\mu\Big(a_{i,j}(r,y,[X_r^{s,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s,\xi,(m)}])\Big)| \leq K\left\{\frac{|z-y|^\eta}{(r-s)^{\frac{\alpha}{2}}}\wedge\frac{1}{(r-s)^{\frac{\alpha-\eta}{2}}}\right\}W_2^\alpha(\mu,\mu'). \qquad (5.83)$$

Moreover, similarly to (5.81), one has

$$\begin{aligned}
&|\partial_v\left[\partial_\mu H_2^{i,j}\left(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m)}])dv',z-y\right)\right](v)\,\widehat{p}_{m+1}(\mu,s,r,t,y,z)|\\
&\leq \frac{K}{(t-r)^2}\int_r^t \max_{i,j}|\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}])](v)|\,dv'\,g(c(t-r),z-y) \qquad (5.84)\\
&\leq \frac{K}{(t-r)(r-s)^{1-\frac{\eta}{2}}}\,g(c(t-r),z-y).
\end{aligned}$$

From the two previous estimates we thus conclude

$$|\mathrm{D}_1| \le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)(r-s)^{1+\frac{\alpha}{2}-\eta}}\right\} W_2^\alpha(\mu,\mu')\, g(c(t-r), z-y).$$

For D_2, we handle $\Delta_\mu\partial_v[\partial_\mu H_2^{i,j}(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m-1)}])dv', z-y)](v)$ like $\Delta_\mu\partial_v[\partial_\mu H_1^i(\int_r^t a(v',z,[X_{v'}^{s,\xi,(m-1)}])dv', z-y)](v)$, that is, from the mean-value theorem and (5.82), one gets

$$\begin{aligned}|\mathrm{D}_2| &\le \frac{K}{(t-r)^{2-\frac{\eta}{2}}}\int_r^t \max_{i,j}|\Delta_\mu\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m-1)}])](v)|dv'\, g(c(t-r),z-y)\\ &\le \frac{K}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\left(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv\right) W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).\end{aligned}$$

To deal with D_3 we employ (5.76) and (5.84). We get

$$\begin{aligned}|\mathrm{D}_3| &\le \frac{K}{(t-r)^{2-\frac{\eta}{2}}(r-s)^{\frac{\alpha-\eta}{2}}}\int_r^t \max_{i,j}|\partial_v[\partial_\mu a_{i,j}(v',z,[X_{v'}^{s,\xi,(m)}])](v)|\,dv'\, g(c(t-r),z-y)\\ &\le K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha}{2}-\eta}}\, g(c(t-r),z-y).\end{aligned}$$

Gathering the previous estimates on D_1, D_2 and D_3, we get

$$\begin{aligned}|\mathrm{D}| \le C&\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\}\left(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v)(v-s)^{\frac{\eta}{2}}\,dv\right)\\ &\times W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).\end{aligned}$$

• **Estimates on** E**:**

For E_1, we proceed as for the previous terms. To be more specific, from (5.80), the mean-value theorem and (5.60) as well as (5.41) (bounding the sum by K), we have

$$|\mathrm{E}_1| \le K\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{\frac{1}{2}}(r-s)^{1+\frac{\alpha}{2}-\eta}}\, g(c(t-r),z-y).$$

For E_2, from (5.83), (5.39) and then (5.79), we get

$$|\mathrm{E}_2| \le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)(r-s)^{1+\frac{\alpha}{2}-\eta}}\right\} W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).$$

For the last term E_3, from (5.73), one obtains

$$\begin{aligned}|\mathrm{E}_3| \le K&\left\{\frac{W_2^\alpha(\mu,\mu')}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\right.\\ &\left.+\frac{1}{(t-r)^{2-\frac{\eta}{2}}}\int_r^t\int\int(|y''-x'|^\eta\wedge 1)|\Delta_\mu\partial_v[\partial_\mu p_m(\mu,s,v',x',y'')](v)|\,dy''\,\mu'(dx')\,dv'\right\}\\ &\times\ g(c(t-r),z-y)\\ \le\ &\frac{K}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\left(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v')(v'-s)^{\frac{\eta}{2}}\,dv'\right) W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).\end{aligned}$$

Gathering the previous estimates, we finally deduce

$$|\mathrm{E}| \le \frac{K}{(t-r)^{1-\frac{\eta}{2}}(r-s)^{1+\frac{\alpha-\eta}{2}}}\left(1+(t-r)^{-1}\int_r^t C_{m,n}(s,v')(v'-s)^{\frac{\eta}{2}}\,dv'\right) W_2^\alpha(\mu,\mu')\, g(c(t-r),z-y).$$

We now collect all the previous estimates on A, B, C, D and E. We finally obtain the following bound

$$\begin{aligned}&|\Delta_\mu\partial_v[\partial_\mu\mathcal{H}_{m+1}(\mu,s,r,t,y,z)](v)|\\ &\le K\left\{\frac{1}{(t-r)(r-s)^{1+\frac{\alpha-\eta}{2}}} \wedge \frac{1}{(t-r)^{1+\frac{\alpha-\eta}{2}}(r-s)}\right\}\left(1+\frac{1}{t-r}\int_r^t C_{m,n}(s,v')(v'-s)^{\frac{\eta}{2}}\,dv'\right) W_2^\alpha(\mu,\mu')g(c(t-r),z-y)\end{aligned}$$

which in turn, after a space-time convolution with p_{m+1}, implies

$$
\begin{aligned}
&|p_{m+1}\otimes\Delta_\mu\partial_v[\partial_\mu\mathcal{H}_{m+1}(\mu,s,t,x,z)](v)|\\
&\quad\leq \frac{K}{(t-s)^{1+\frac{\alpha-\eta}{2}}}\left(B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}\right)+\sum_{k=1}^{m-1}C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right)\\
&\qquad\times W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x)
\end{aligned}
$$

where we again separate the time integral according to the two disjoint intervals: $[s,\frac{t+s}{2}]$ and $(\frac{t+s}{2},t]$ in order to balance the time singulairity. From standard computations, we deduce that the series $\sum_{k\geq 0}(p_{m+1}\otimes\Delta_\mu\partial_v\partial_\mu\mathcal{H}_{m+1})\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(v)$ converges absolutely and uniformly. Moreover, there exist positive constants $K:=K(T,a,b)$, $c:=c(\lambda)$ such that for any $\alpha\in[0,\eta)$

$$
\begin{aligned}
&\sum_{k\geq 0}\left|\left(p_{m+1}\otimes\Delta_\mu\partial_v\partial_\mu\mathcal{H}_{m+1}\right)\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(v)\right|\\
&\quad\leq \frac{K}{(t-s)^{1+\frac{\alpha-\eta}{2}}}\left\{B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}\right)+\sum_{k=1}^{m}C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}}\right\}\\
&\qquad\times W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x).
\end{aligned}
\tag{5.85}
$$

From (5.63) and (5.12) (at step $m+1$), separating the computations into the two disjoint intervals $[s,\frac{t+s}{2}]$ and $[\frac{t+s}{2},t]$ as in (5.67) and (5.68), we get

$$
\forall\alpha\in[0,\eta),\quad |(\partial_v\partial_\mu p_{m+1}\otimes\Delta_\mu\mathcal{H}_{m+1})(\mu,s,t,x,z)(v)|\leq\frac{K}{(t-s)^{1+\frac{\alpha}{2}-\eta}}W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x)
$$

so that

$$
\forall\alpha\in[0,\eta),\quad \sum_{k\geq 0}|(\partial_v\partial_\mu p_{m+1}\otimes\Delta_\mu\mathcal{H}_{m+1})\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(v)|\leq\frac{K}{(t-s)^{1+\frac{\alpha}{2}-\eta}}W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x).
\tag{5.86}
$$

Similarly, from the estimates (5.59), (5.47) (bounding $C_{m,n}$ by a constant K), separating the time integral into two disjoint intervals as previously done, we get

$$
\forall\alpha\in[0,\eta),\quad |\Delta_\mu p_{m+1}\otimes\partial_v\partial_\mu\mathcal{H}_{m+1}(\mu,s,t,x,z)(v)|\leq\frac{K}{(t-s)^{1+\frac{\alpha}{2}-\eta}}W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x)
$$

which in turn implies

$$
\forall\alpha\in[0,\eta),\quad \sum_{k\geq 0}|(\Delta_\mu p_{m+1}\otimes\partial_v\partial_\mu\mathcal{H}_{m+1})\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(v)|\leq\frac{K}{(t-s)^{1+\frac{\alpha}{2}-\eta}}W_2^\alpha(\mu,\mu')\,g(c(t-s),z-x).
\tag{5.87}
$$

If we differentiate with respect to the measure argument (and then with respect to the variable v) the relation $p_{m+1}=\widehat{p}_{m+1}+p_{m+1}\otimes\mathcal{H}_{m+1}$, we obtain $\partial_v\partial_\mu p_{m+1}=\partial_v\partial_\mu\widehat{p}_{m+1}+p_{m+1}\otimes\partial_v\partial_\mu\mathcal{H}_{m+1}+\partial_v\partial_\mu p_{m+1}\otimes\mathcal{H}_{m+1}$ so that

$$
\begin{aligned}
\Delta_\mu\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)&=\Delta_\mu\partial_v[\partial_\mu\widehat{p}_{m+1}(\mu,s,t,x,z)](v)+p_{m+1}\otimes\Delta_\mu\partial_v[\partial_\mu\mathcal{H}_{m+1}(\mu,s,t,x,z)](v)\\
&\quad+\Delta_\mu p_{m+1}\otimes\partial_v\partial_\mu\mathcal{H}_{m+1}(\mu,s,t,x,z)(v)+\partial_v\partial_\mu p_{m+1}\otimes\Delta_\mu\mathcal{H}_{m+1}(\mu,s,t,x,z)(v)\\
&\quad+\Delta_\mu\partial_v\partial_\mu p_{m+1}\otimes\mathcal{H}_{m+1}(\mu,s,t,x,z)(v).
\end{aligned}
$$

Iterating the previous relation, we obtain the following representation

$$
\begin{aligned}
\Delta_\mu\partial_v[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)=\sum_{k\geq 0}\Big[&\Delta_\mu\partial_v\partial_\mu\widehat{p}_{m+1}+p_{m+1}\otimes\Delta_\mu\partial_v\partial_\mu\mathcal{H}_{m+1}\\
&+\Delta_\mu p_{m+1}\otimes\partial_v\partial_\mu\mathcal{H}_{m+1}+\partial_v\partial_\mu p_{m+1}\otimes\Delta_\mu\mathcal{H}_{m+1}\Big]\otimes\mathcal{H}_{m+1}^{(k)}(\mu,s,t,x,z)(v).
\end{aligned}
$$

Gathering the estimates (5.74), (5.85), (5.86) and (5.87), we deduce that the above series converges absolutely and satifies

$$\begin{aligned}\forall \alpha \in [0,\eta), \quad & |\Delta_\mu \partial_v [\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)| \\ & \leq \frac{K}{(t-s)^{1+\frac{\alpha}{2}-\eta}} W_2^\alpha(\mu,\mu')\, g(c(t-s),z-x) \\ & \quad + \frac{K}{(t-s)^{1+\frac{\alpha-\eta}{2}}} \left\{ B\left(\frac{\eta}{2},\frac{\alpha-\eta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\} \\ & \quad \times W_2^\alpha(\mu,\mu')\, g(c(t-s),z-x) \\ & \leq \frac{K}{(t-s)^{1+\frac{\alpha-\eta}{2}}} \left\{ B\left(\frac{\eta}{2},\frac{\alpha-\eta}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\} \\ & \quad \times W_2^\alpha(\mu,\mu') g(c(t-s),z-x)\end{aligned}$$

so that

$$u_{m+1}^n(s,t) \leq \frac{K}{(t-s)^{1+\frac{\alpha}{2}-\eta}} \left\{ B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\}$$

and similarly,

$$v_{m+1}^n(s,t) \leq \frac{K}{(t-s)^{1+\frac{\alpha-\eta}{2}}} \left\{ B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}\right) + \sum_{k=1}^m C^k \prod_{i=1}^{k+1} B\left(\frac{\eta}{2},\frac{\eta-\alpha}{2}+(i-1)\frac{\eta}{2}\right)(t-s)^{k\frac{\eta}{2}} \right\}.$$

Since the constant K does not depend either on the constant C appearing in the definition of $C_{m,n}(s,t)$ or m, one may change C once for all and derive the induction hypothesis at step $m+1$ for u_m^n and v_m^n. This completes the proof of (5.17).

We now prove the estimates (5.18) and (5.19). Since the proofs are rather long, technical and use similar arguments as those employed before, we will limit ourself to (5.19) and will omit some technical details. The proof of (5.18) follows from the relation (5.23) and similar arguments as those developed below. We remark that if $|s_1-s_2| \geq t-s_1\vee s_2$, the estimate (5.19) follows directly from (5.12). We thus assume that $|s_1-s_2| \leq t-s_1\vee s_2$ for the rest of the proof. To make the notations simpler, for any fixed $(s_1,s_2)\in[0,t)^2$, we write $\Delta_s f(s) = f(s_1\vee s_2) - f(s_1\wedge s_2)$ where f is a function defined on $[0,t)$. In particular, $\Delta_s p_m(\mu,s,t,x,z) = p_m(\mu,s_1\vee s_2,t,x,z) - p_m(\mu,s_1\wedge s_2,t,x,z)$. We first claim

$$\begin{aligned}\forall \beta\in[0,1], \forall m\geq 1, \quad & |\Delta_s p_m(\mu,s,t,x,z)| \\ & \leq K\left\{ \frac{|s_1-s_2|^\beta}{(t-s_1)^\beta} g(c(t-s_1),z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^\beta} g(c(t-s_2),z-x)\right\}. \end{aligned} \tag{5.88}$$

In order to prove the above statement, one has to consider the two cases $|s_1-s_2| \geq t-s_1\vee s_2$ and $|s_1-s_2| \leq t-s_1\vee s_2$. In the first case, it directly follows from (5.6) with $n=0$, while in the second case, it follows from the mean-value theorem, (5.13) and the inequality $(t-s_1\vee s_2)^{-1} \leq 2(t-s_1\wedge s_2)^{-1}$.

We now start from the representation in infinite series (5.40) and write the following decomposition

$$\begin{aligned}\Delta_s \partial_v^n [\partial_\mu p_{m+1}(\mu,s,t,x,z)](v) &= \Delta_s \partial_v^n [\partial_\mu \widehat{p}_{m+1}(\mu,s,t,x,z)](v) + \Delta_s(p_{m+1}\otimes \partial_v^n[\partial_\mu \mathcal{H}_{m+1}])(\mu,s,t,x,z)(v) \\ &\quad + \Delta_s(\partial_v^n[\partial_\mu \widehat{p}_{m+1}]\otimes \Phi_{m+1})(\mu,s,t,x,z)(v) \\ &\quad + \Delta_s((p_{m+1}\otimes \partial_y^n[\partial_\mu \mathcal{H}_{m+1}])\otimes \Phi_{m+1})(\mu,s,t,x,z)(v).\end{aligned} \tag{5.89}$$

We investigate the first term appearing in the right-hand side of the above identity and make use of the following decomposition

$$\Delta_s \partial_v^n [\partial_\mu \widehat{p}_{m+1}(\mu,s,t,x,z)](v) = \mathrm{I}(v) + \mathrm{II}(v) + \mathrm{III}(v) + \mathrm{IV}(v),$$

with

$$\begin{aligned}
\mathrm{I}(v) &:= \left\{ Df_{z-x}\left(\int_{s_1\vee s_2}^t a(r,z,[X_r^{s_1\vee s_2,\xi,(m)}])dr\right) - Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])dr\right)\right\} \\
&\cdot\int_{s_1\vee s_2}^t \left\{ \int (\widetilde{a}(r,z,y',[X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde{a}(r,z,v,[X_r^{s_1\vee s_2,\xi,(m)}]))\partial_x^{1+n} p_m(\mu,s_1\vee s_2,r,v,y')dy' \right. \\
&\left. + \int\int (\widetilde{a}(r,z,y',[X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde{a}(r,z,x',[X_r^{s_1\vee s_2,\xi,(m)}]))\partial_v^n\partial_\mu[p_m(\mu,s_1\vee s_2,r,x',y')](v)\,dy'\mu(dx')\right\} dr, \\
\mathrm{II}(v) &:= Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])dr\right) \\
&\cdot\int_{s_1\vee s_2}^t \left\{ \int (\widetilde{a}(r,z,y',[X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde{a}(r,z,y',[X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_x^{1+n} p_m(\mu,s_1\vee s_2,r,v,y')\,dy' \right. \\
&+ \int (\widetilde{a}(r,z,y',[X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde{a}(r,z,v,[X_r^{s_1\vee s_2,\xi,(m)}])) \\
&\left. .(\partial_x^{1+n} p_m(\mu,s_1\vee s_2,r,v,y') - \partial_x^{1+n} p_m(\mu,s_1\wedge s_2,r,v,y'))dy'\right\} dr, \\
\mathrm{III}(v) &:= Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r,y,[X_r^{s_1\wedge s_2,\xi,(m)}])dr\right) \\
&\cdot\int_{s_1\vee s_2}^t \left\{ \int\int (\widetilde{a}(r,z,y',[X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde{a}(r,z,y',[X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu,s_1\vee s_2,r,x',y')](v)dy'\mu(dx') \right. \\
&+ \int\int (\widetilde{a}(r,z,y',[X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde{a}(r,z,x',[X_r^{s_1\wedge s_2,\xi,(m)}])) \\
&\left. .\left(\partial_v^n[\partial_\mu p_m(\mu,s_1\vee s_2,r,x',y')](v) - \partial_v^n[\partial_\mu p_m(\mu',s_1\wedge s_2,r,x',y')](v)\right) dy'\mu(dx')\right\} dr, \\
\mathrm{IV}(v) &:= -Df_{z-x}\left(\int_{s_1\wedge s_2}^t a(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])dr\right) \\
&\cdot\int_{s_1\wedge s_2}^{s_1\vee s_2} \left\{ \int (\widetilde{a}(r,z,y',[X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde{a}(r,z,v,[X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_x^{1+n} p_m(\mu,s_1\wedge s_2,r,v,y')dy' \right. \\
&\left. + \int\int (\widetilde{a}(r,z,y',[X_r^{s_1\wedge s_2,\xi,(m)}]) - \widetilde{a}(r,z,x',[X_r^{s_1\wedge s_2,\xi,(m)}]))\partial_v^n[\partial_\mu p_m(\mu,s_1\wedge s_2,r,x',y')](v)dy'\mu(dx')\right\} dr.
\end{aligned}$$

From the mean-value theorem, (**HR**)(ii) with (5.88), (5.12) and (5.6), we get

$$\begin{aligned}
\forall\beta\in[0,1],\ |\mathrm{I}(v)| &\le \frac{K}{(t-s_1\vee s_2)^2}\left\{|s_1-s_2| + \int_{s_1\vee s_2}^t \frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}}\,dr\right\}(t-s_1\vee s_2)^{\frac{1-n+\eta}{2}} \\
&\quad\times g(c(t-s_1\wedge s_2),z-x) \\
&\le K\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\,g(c(t-s_1\wedge s_2),z-x)
\end{aligned}$$

where we used the inequality $|s_1-s_2|\le t-s_1\vee s_2$ for the last line. From $(\mathbf{CS}_+)_2$ with (5.88) and the η-Hölder regularity of $z\mapsto\widetilde{A}(t,x,z,z',\nu,\mu)$, one gets

$$|\widetilde{a}(r,z,y',[X_r^{s_1\vee s_2,\xi,(m)}]) - \widetilde{a}(r,z,y',[X_r^{s_1\wedge s_2,\xi,(m)}])| \le K\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}} \tag{5.90}$$

which, together with (5.6), (5.18) both with $n=1,2$ and the η-Hölder regularity of $y\mapsto\widetilde{a}(r,z,y,\mu)$, imply

$$\begin{aligned}
|\mathrm{II}(v)| &\le \frac{K}{t-s_1\vee s_2}\int_{s_1\vee s_2}^t \frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\,dr\,g(c(t-s_1\wedge s_2),z-x) \\
&\le K\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\,g(c(t-s_1\wedge s_2),z-x)
\end{aligned}$$

where $\beta \in [0, \frac{1+\eta}{2})$ for $n = 0$ and $\beta \in [0, \frac{\eta}{2})$ for $n = 1$. From (5.90), (5.12) and $(\mathbf{CS}_+)_1$, we get

$$|\mathrm{III}(v)| \le K\Big\{ \frac{|s_1 - s_2|^\beta}{(t - s_1 \vee s_2)^{\frac{1+n}{2}+\beta-\eta}} + \frac{1}{t - s_1 \wedge s_2} \int_{s_1\vee s_2}^t \int\int (|y'' - x'|^\eta \wedge 1) \\ |\partial_v^n[\partial_\mu p_m(\mu, s_1 \vee s_2, r, x', y'')](v) - \partial_v^n[\partial_\mu p_m(\mu, s_1 \wedge s_2, r, x', y'')](v)| dr\, dy'' \mu(dx') \Big\} g(c(t - s_1 \wedge s_2), z - x)$$

where $\beta \in [0, (\frac{1}{2} + \eta) \wedge 1)$ for $n = 0$ and $\beta \in [0, \eta)$ for $n = 1$. From (5.6), (5.12) and the η-Hölder regularity of $y \mapsto \widetilde{a}(r, z, y, \mu)$, we finally obtain

$$|\mathrm{IV}(v)| \le K \frac{|s_1 - s_2|^\beta}{(t - s_1 \vee s_2)^{\frac{1+n-\eta}{2}+\beta}} g(c(t - s_1 \wedge s_2), z - x)$$

where $\beta \in [0, \frac{1+\eta}{2})$ for $n = 0$ and $\beta \in [0, \frac{\eta}{2})$ for $n = 1$. Collecting the above estimates, we finally obtain

$$\begin{aligned} &|\Delta_s \partial_v^n[\partial_\mu \widehat{p}_{m+1}(\mu, s, t, x, z)](v)| \\ &\le K\Big\{ \frac{|s_1 - s_2|^\beta}{(t - s_1 \vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{1}{t - s_1 \wedge s_2} \int_{s_1\vee s_2}^t \int\int (|y'' - x'|^\eta \wedge 1) \\ &\quad \times |\partial_v^n[\partial_\mu p_m(\mu, s_1 \vee s_2, r, x', y'')](v) - \partial_v^n[\partial_\mu p_m(\mu, s_1 \wedge s_2, r, x', y'')](v)|\, dy'' \mu(dx') dr \Big\} \\ &\quad \times g(c(t - s_1 \wedge s_2), z - x). \end{aligned} \tag{5.91}$$

As we already did before we introduce the quantities

$$u_m^n(s_1, s_2, t) := \sup_{(s_1,s_2,v)\in[0,t)^2\times\mathbb{R}^d, s_1\neq s_2} \int\int (|y - x|^\eta \wedge 1) \frac{|\Delta_s \partial_v^n[\partial_\mu p_m(\mu, s, t, x, y)](v)|}{|s_1 - s_2|^\beta} dy\, \mu(dx)$$
$$v_m^n(s_1, s_2, t) := \sup_{(s_1,s_2,v)\in[0,t)^2\times\mathbb{R}^d, s_1\neq s_2} \int\int \frac{|\Delta_s \partial_v^n[\partial_\mu p_m(\mu, s, t, x, y)](v)|}{|s_1 - s_2|^\beta} dy\, \mu(dx)$$

for any fixed $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and any fixed $\beta \in [0, \eta/2)$ for $n = 0, 1$. We prove by induction the following key inequalities:

$$u_m^n(s_1, s_2, t) \le C_{m,n}(s_1 \vee s_2, t)(t - s_1 \vee s_2)^{-\frac{(1+n)}{2}-\beta+\eta},$$
$$v_m^n(s_1, s_2, t) \le C_{m,n}(s_1 \vee s_2, t)(t - s_1 \vee s_2)^{-\frac{(1+n)}{2}-\beta+\frac{\eta}{2}},$$

with $C_{m,n}(s, t) := \sum_{k=1}^m C^k \prod_{i=1}^k B\left(\frac{\eta}{2}, \frac{1-n+\eta}{2} - \beta + (i-1)\frac{\eta}{2}\right)(t - s)^{(k-1)\frac{\eta}{2}}$. The result for $m = 1$ being straightforward, we assume that it holds at step m. With the above induction hypothesis applied to (5.91), we get

$$\begin{aligned} |\Delta_s \partial_v^n[\partial_\mu \widehat{p}_{m+1}(\mu, s, t, x, z)](v)| &\le K\Big\{ \frac{|s_1 - s_2|^\beta}{(t - s_1 \vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{1}{t - s_1 \wedge s_2} \int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1 \vee s_2, r)}{(r - s_1 \vee s_2)^{\frac{1+n}{2}+\beta-\eta}} dr \Big\} \\ &\quad \times g(c(t - s_1 \wedge s_2), z - x). \end{aligned} \tag{5.92}$$

From $\Phi_{m+1}(\mu, s, r, t, x, y) = \mathcal{H}_{m+1}(\mu, s, r, t, x, y) + \mathcal{H}_{m+1} \otimes \Phi_{m+1}(\mu, s, r, t, x, y)$, we obtain the following relation

$$\begin{aligned} \Delta_s \Phi_{m+1}(\mu, s, r, t, x, y) &= \Delta_s \mathcal{H}_{m+1}(\mu, s, r, t, x, y) \\ &\quad + \int_r^t \int \Delta_s \mathcal{H}_{m+1}(\mu, s, r, v, x, z) \Phi_{m+1}(\mu, s_1 \vee s_2, v, t, z, y)\, dv\, dy \\ &\quad + \int_r^t \int \mathcal{H}_{m+1}(\mu, s_1 \wedge s_2, r, v, x, z) \Delta_s \Phi_{m+1}(\mu, s, v, t, z, y)\, dv\, dy. \end{aligned} \tag{5.93}$$

We claim

$$|\Delta_s \mathcal{H}_{m+1}(\mu, s, r, t, x, y)| \le K \left\{ \frac{|s_1 - s_2|^\beta}{(t - r)^{1-\frac{\eta}{2}}(r - s_1 \vee s_2)^\beta} \wedge \frac{|s_1 - s_2|^\beta}{(t - r)(r - s_1 \vee s_2)^{\beta-\frac{\eta}{2}}} \right\} g(c(t - r), y - x). \tag{5.94}$$

In order to prove the above statement, we use a similar decomposition as the one employed for $\Delta_\mu \mathcal{H}_{m+1}(\mu, s, r, t, x, y)$, namely

$$\Delta_s \mathcal{H}_{m+1}(\mu, s, r, t, x, y) = \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV} + \mathrm{V}$$

with

$$\mathrm{I} := -\sum_{i=1}^{d} \Delta_s b_i(r,x,[X_r^{s,\xi,(m)}]) H_1^i \left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\, dv, z-x \right) \widehat{p}_{m+1}(\mu, s_1\vee s_2, r, t, x, z),$$

$$\mathrm{II} := -\sum_{i=1}^{d} b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}]) \Delta_s H_1^i \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\, dv, z-x \right) \widehat{p}_{m+1}(\mu, s_1\vee s_2, r, t, x, z)$$

$$\mathrm{III} := \frac{1}{2}\sum_{i,j=1}^{d} \Delta_s [a_{i,j}(r,x,[X_r^{s,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s,\xi,(m)}])] H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\, dv, z-x \right)$$
$$\times \widehat{p}_{m+1}(\mu, s_1\vee s_2, r, t, x, z),$$

$$\mathrm{IV} := \frac{1}{2}\sum_{i,j=1}^{d} [a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])] \Delta_s H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\, dv, z-x \right)$$
$$\times \widehat{p}_{m+1}(\mu, s_1\vee s_2, r, t, x, z),$$

$$\mathrm{V} := -\sum_{i=1}^{d} b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}]) H_1^i \left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\, dv, z-x \right) \Delta_s \widehat{p}_{m+1}(\mu, s, r, t, x, z),$$
$$+ \frac{1}{2}\sum_{i,j=1}^{d} [a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])] H_2^{i,j} \left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\, dv, z-x \right)$$
$$\times \Delta_s \widehat{p}_{m+1}(\mu, s, r, t, x, z).$$

From (**HR**)(ii), (5.88) and (5.13), following similar arguments as those employed for $\Delta_\mu \mathcal{H}_{m+1}(\mu,s,r,t,x,y)$, the following estimates hold: for all $\beta \in [0,1]$,

$$|\mathrm{I}| \le K \frac{|s_1-s_2|^\beta}{(t-r)^{\frac{1}{2}}(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}}\, g(c(t-r), y-x),$$

$$|\mathrm{II}| \le K \frac{|s_1-s_2|^\beta}{(t-r)^{\frac{1}{2}}(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}}\, g(c(t-r), y-x),$$

$$|\mathrm{III}| \le K \left\{ \frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\beta}} \wedge \frac{1}{(t-r)(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}} \right\} |s_1-s_2|^\beta\, g(c(t-r), y-x),$$

$$|\mathrm{IV}| \le K \frac{|s_1-s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}}\, g(c(t-r), y-x),$$

$$|\mathrm{V}| \le K \frac{|s_1-s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}}\, g(c(t-r), y-x).$$

We only prove the estimates on I and III. The estimates on II, IV and V are obtained by following similar lines of reasonings and the remaining technical details are omitted. From (**HR**)(ii) and (5.88), similarly to (5.90) with B instead of the map $\widetilde{a}$, we get

$$|\Delta_s b_i(r,x,[X_r^{s,\xi,(m)}])| \le K \frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}} \tag{5.95}$$

which in turn directly yields the announced estimates on I. In order to deal with III, we consider the two disjoint cases: $|s_1-s_2| \ge (r-s_1\vee s_2)$ and $|s_1-s_2| \le (r-s_1\vee s_2)$. In the first case, from the η-Hölder regularity of $x \mapsto a(t,x,\mu)$ and the space-time inequality (1.4), we directly obtain

$$\forall \beta \in [0,1],\ |\mathrm{III}| \le K \frac{1}{(t-r)^{1-\frac{\eta}{2}}} g(c(t-r), z-x) \le K \frac{|s_1-s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\beta}} g(c(t-r), z-x).$$

In order to obtain the other part of the estimate, we combine (**HR**)(ii) with (5.88), similarly to (5.95), we get $|\Delta_s a_{i,j}(r,x,[X_r^{s,\xi,(m)}])| \le K|s_1-s_2|^\beta (r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}$ which in turn readily implies

$$\forall \beta \in [0,1],\ |\mathrm{III}| \le K \frac{|s_1-s_2|^\beta}{(t-r)(r-s_1\vee s_2)^{\beta-\frac{\eta}{2}}} g(c(t-r), z-x). \tag{5.96}$$

We thus derive the announced estimate on III in the case $|s_1-s_2| \ge (r-s_1\vee s_2)$. We importantly observe that (5.96) is still valid in the case $|s_1-s_2| \le (r-s_1\vee s_2)$. Now assume that $|s_1-s_2| \le (r-s_1\vee s_2)$.

Then, from the mean-value theorem and Fubini's theorem, one gets

$$\begin{aligned}h(x) &:= a_{i,j}(r,x,[X_r^{s_1\vee s_2,\xi,(m)}]) - a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])\\ &= \int_0^1 \int\int \widetilde{a}_{i,j}(r,x,y'',[X_r^{\lambda s_1\vee s_2+(1-\lambda)s_1\wedge s_2,\xi,(m)}])\\ &\quad\times \partial_s p_m(\mu,\lambda s_1\vee s_2+(1-\lambda)s_1\wedge s_2,r,x',y'')(s_1\vee s_2-s_1\wedge s_2)\,dy''\mu(dx').\end{aligned}$$

The previous identity together with the η-Hölder regularity of $x\mapsto \widetilde{a}_{i,j}(r,x,y,\mu)$ and (5.13) yield

$$\forall \beta\in[0,1],\ |h(x)-h(y)|\le K(|y-x|^\eta\wedge 1)\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^\beta}$$

which, combined with the space-time inequality (1.4) directly imply

$$\forall \beta\in[0,1],\ |\mathrm{III}|\le K\frac{|s_1-s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^\beta}g(c(t-r),z-x).$$

This last estimate concludes the proof of the announced result on III. Gathering the previous estimates allow to conclude that (5.94) holds. With the previous result, we derive an estimate for the second term appearing in the right-hand side of (5.93). From (5.94), one gets $|\Delta_s\mathcal{H}_{m+1}(\mu,s,r,v,x,z)|\le K|s_1-s_2|^\beta(v-r)^{-1+\frac{\eta}{2}}(r-s_1\vee s_2)^{-\beta}g(c(v-r),z-x)$, so that, after some standard computations, we get

$$\Big|\int_r^t\int \Delta_s\mathcal{H}_{m+1}(\mu,s,r,v,x,z)\Phi_{m+1}(\mu,s_1\vee s_2,v,t,z,y)\,dv\,dy\Big| \le K\frac{|s_1-s_2|^\beta}{(t-r)^{1-\eta}(r-s_1\vee s_2)^\beta}\,g(c(t-r),y-x)$$

which in turn, from the identity (5.93), the estimate (5.94) and a direct induction argument, yield

$$\forall\beta\in[0,1],\quad |\Delta_s\Phi_{m+1}(\mu,s,r,t,x,y)|\le K\frac{|s_1-s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^\beta}\,g(c(t-r),y-x). \tag{5.97}$$

From (5.38) we obtain the following decomposition

$$\Delta_s\partial_y^n[\partial_\mu\mathcal{H}_{m+1}(\mu,s,r,t,x,z)](y)=\mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}+\mathrm{V}$$

with

$$\begin{aligned}\mathrm{I} &= -\sum_{i=1}^d \Delta_s\partial_y^n[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])](y)H_1^i\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv,z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\ &\quad-\sum_{i=1}^d \partial_y^n[\partial_\mu b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])](y)\Delta_s H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv,z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\ &\quad-\sum_{i=1}^d \partial_y^n[\partial_\mu b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])](y)H_1^i\left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\,dv,z-x\right)\Delta_s\widehat{p}_{m+1}(\mu,s,r,t,x,z)\\ &=: \mathrm{I}_1+\mathrm{I}_2+\mathrm{I}_3,\end{aligned}$$

$$\begin{aligned}\mathrm{II} &= -\sum_{i=1}^d \Delta_s b_i(r,x,[X_r^{s,\xi,(m)}])\partial_y^n[\partial_\mu H_1^i\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv,z-x\right)](y)\,\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\ &\quad-\sum_{i=1}^d b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])\Delta_s\partial_y^n[\partial_\mu H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv,z-x\right)](y)\,\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\ &\quad-\sum_{i=1}^d b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])\partial_y^n[\partial_\mu H_1^i\left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\,dv,z-x\right)](y)\,\Delta_s\widehat{p}_{m+1}(\mu,s,r,t,x,z)\\ &=: \mathrm{II}_1+\mathrm{II}_2+\mathrm{II}_3,\end{aligned}$$

$$
\begin{aligned}
\mathrm{III} &= \frac{1}{2}\sum_{i,j=1}^{d}\Delta_s\partial_y^n[\partial_\mu[a_{i,j}(r,x,[X_r^{s,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s,\xi,(m)}])]](y)H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)\\
&\quad\times\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\
&\quad+\frac{1}{2}\sum_{i,j=1}^{d}\partial_y^n[\partial_\mu[a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])]](y)\Delta_s H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)\\
&\qquad\times\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\
&\quad+\frac{1}{2}\sum_{i,j=1}^{d}\partial_y^n[\partial_\mu[a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])]](y)H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\,dv, z-x\right)\\
&\qquad\times\Delta_s\widehat{p}_{m+1}(\mu,s,r,t,x,z)\\
&=:\mathrm{III}_1+\mathrm{III}_2+\mathrm{III}_3,
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{IV} &= \frac{1}{2}\sum_{i,j=1}^{d}\Delta_s[a_{i,j}(r,x,[X_r^{s,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s,\xi,(m)}])]\partial_y^n[\partial_\mu H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)](y)\\
&\quad\times\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\
&\quad+\frac{1}{2}\sum_{i,j=1}^{d}[a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])]\Delta_s\partial_y^n[\partial_\mu H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)](y)\\
&\qquad\times\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)\\
&\quad+\frac{1}{2}\sum_{i,j=1}^{d}[a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])]\partial_y^n[\partial_\mu H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\,dv, z-x\right)](y)\\
&\qquad\times\Delta_s\widehat{p}_{m+1}(\mu,s,r,t,x,z)\\
&=:\mathrm{IV}_1+\mathrm{IV}_2+\mathrm{IV}_3,
\end{aligned}
$$

and finally

$$
\begin{aligned}
\mathrm{V} &= -\sum_{i=1}^{d}\Delta_s[b_i(r,x,[X_r^{s,\xi,(m)}])H_1^i\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)]\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)](y)\\
&\quad+\frac{1}{2}\sum_{i,j=1}^{d}\Delta_s[a_{i,j}(r,x,[X_r^{s,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s,\xi,(m)}])H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)]\\
&\qquad\times\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)](y)\\
&\qquad+\left\{-\sum_{i=1}^{d}b_i(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])H_1^i\left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\,dv, z-x\right)\right.\\
&\qquad\left.+\frac{1}{2}\sum_{i,j=1}^{d}[a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}])-a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])]H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s_1\wedge s_2,\xi,(m)}])\,dv, z-x\right)\right\}\\
&\qquad\times\Delta_s\partial_y^n[\partial_\mu\widehat{p}_{m+1}(\mu,s,r,t,x,z)](y)\\
&=:\mathrm{V}_1+\mathrm{V}_2+\mathrm{V}_3.
\end{aligned}
$$

From similar arguments as those employed for the proofs of the estimates appearing in the decomposition of $\Delta_\mu\partial_v[\partial_\mu\mathcal{H}_{m+1}(\mu,s,t,x,z)](v)$, we obtain the following bounds:

$$
|\mathrm{I}|\le K\left\{\frac{1}{(t-r)^{\frac{1}{2}}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}+\frac{1}{(t-r)^{\frac{1}{2}}}u_m^n(s_1,s_2,r)\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x),
$$

$$
|\mathrm{II}|\le K\left\{\frac{1}{(t-r)^{\frac{1}{2}}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}+\frac{1}{(t-r)^{\frac{3}{2}}}\int_r^t u_m^n(s_1,s_2,v)\,dv\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x),
$$

$$|\mathrm{III}| \le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r-s_1\vee s_2)^{\frac{1+n}{2}}} \wedge \frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{v_m^n(s_1,s_2,r)}{(t-r)^{1-\frac{\eta}{2}}} \wedge \frac{u_m^n(s_1,s_2,r)}{t-r}\right\} \times |s_1-s_2|^\beta\, g(c(t-r),z-x),$$

$$|\mathrm{IV}| \le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{1}{(t-r)^{2-\frac{\eta}{2}}}\int_r^t u_m^n(s_1,s_2,v)\,dv\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x),$$

and

$$|\mathrm{V}| \le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{1}{(t-r)^{2-\frac{\eta}{2}}}\int_r^t u_m^n(s_1,s_2,v)\,dv\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x)$$

where $\beta\in[0,1]$ except for III where $\beta\in[0,\eta/2)$. We only prove the estimates on I and III. The estimates on II, IV and V follow from similar lines of reasonings and technical details are omitted. In order to prove the announced estimate on I, we proceed as follows. For I_1, we use the following decomposition

$$\Delta_s\partial_y^n[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])](y) = \mathrm{I}+\mathrm{II}+\mathrm{III}+\mathrm{IV}+\mathrm{V}$$

with

$$\mathrm{I} := \int \Delta_s\widetilde{b}_i(r,x,y'',[X_r^{s,\xi,(m)}])\,\partial_x^{1+n}p_m(\mu,s_1\vee s_2,r,y,y'')\,dy'',$$

$$\mathrm{II} := \int (\widetilde{b}_i(r,x,y'',[X_r^{s_1\wedge s_2,\xi',(m)}]) - \widetilde{b}_i(r,x,y,[X_r^{s_1\wedge s_2,\xi',(m)}])\,\Delta_s\partial_x^{1+n}p_m(\mu,s,r,y,y'')\,dy'',$$

$$\mathrm{III} := \int\int \Delta_s\widetilde{b}_i(r,x,y'',[X_r^{s,\xi,(m)}])\,\partial_y^n[\partial_\mu p_m(\mu,s_1\vee s_2,r,x',y'')](y)\,dy''\,\mu(dx'),$$

$$\mathrm{IV} := \int\int (\widetilde{b}_i(r,x,y'',[X_r^{s_1\wedge s_2,\xi',(m)}]) - \widetilde{b}_i(r,x,x',[X_r^{s_1\wedge s_2,\xi',(m)}])).\Delta_s\partial_y^n[\partial_\mu p_m(\mu,s,r,x',y'')](y)\,dy''\,\mu(dx').$$

From (**HR**)(ii), (5.88), similarly to (5.95) with the map $\widetilde{b}_i$ instead of b_i, and (5.6), we get

$$|\mathrm{I}| \le K\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}.$$

For II, we use (5.18) and the η-Hölder regularity of $z\mapsto\widetilde{b}_i(r,x,z,\mu)$

$$|\mathrm{II}| \le K\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}.$$

For III, similarly to I, using (5.12) instead of (5.6), we get

$$|\mathrm{III}| \le K\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}.$$

Finally, for the last term, from the η-Hölder regularity of $x'\mapsto\widetilde{b}_i(r,y,x',\mu)$, one has

$$|\mathrm{IV}| \le K\int\int(|y''-x'|^\eta\wedge 1)|\Delta_s\partial_y^n[\partial_\mu p_m(\mu,s,r,x',y'')](y)|\,dy''\,\mu'(dx').$$

Gathering the previous estimates and using the induction hypothesis, we finally obtain

$$|\Delta_s\partial_y^n[\partial_\mu b_i(r,x,[X_r^{s,\xi,(m)}])](y)| \le K\Big(\frac{1}{(r-s_1\vee s_2)^{\frac{1+n+\beta-\eta}{2}}} + u_m^n(s_1,s_2,r)\Big)\,|s_1-s_2|^\beta$$

so that

$$|\mathrm{I}_1| \le K\left\{\frac{1}{(t-r)^{\frac12}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{1}{(t-r)^{\frac12}}u_m^n(s_1,s_2,r)\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x).$$

From the mean-value theorem, (**HR**) and (5.88), we get

$$|\Delta_s H_1^i\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)| \le K|s_1-s_2|^\beta(t-r)^{-\frac12}(r-s_1\vee s_2)^{-\beta+\frac{\eta}{2}}g(c(t-r),z-x)$$

which with (5.30) imply

$$|\mathrm{I}_2| \leq K \frac{|s_1 - s_2|^\beta}{(t-r)^{\frac{1}{2}}(r - s_1 \vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\, g(c(t-r), z-x).$$

Finally, from (**HR**), (5.88), similarly to (5.95) with the map a instead of b_i, and the mean value theorem, one gets $|\Delta_s \widehat{p}_{m+1}(\mu, s, r, t, x, z)| \leq K|s_1 - s_2|^\beta (r - s_1 \vee s_2)^{-\beta+\frac{\eta}{2}} g(c(t-r), z-x)$ which with (5.30) imply

$$|\mathrm{I}_3| \leq K \frac{|s_1 - s_2|^\beta}{(t-r)^{\frac{1}{2}}(r - s_1 \vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\, g(c(t-r), z-x).$$

Gathering the previous estimates, we obtain the announced estimate on I. In order to deal with III_1, we make use of the decomposition

$$\Delta_s \partial_y^n [\partial_\mu [a_{i,j}(r, x, [X_r^{s,\xi,(m)}]) - a_{i,j}(r, z, [X_r^{s,\xi,(m)}])]](y) = \mathrm{I}_{i,j} + \mathrm{II}_{i,j} + \mathrm{III}_{i,j} + \mathrm{IV}_{i,j}$$

with

$$\begin{aligned}
\mathrm{I}_{i,j} &:= \int \Delta_s [\widetilde{a}_{i,j}(r, x, y'', [X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r, z, y'', [X_r^{s,\xi,(m)}])]\, \partial_x^{1+n} p_m(\mu, s_1 \vee s_2, r, y, y'')\, dy'', \\
\mathrm{II}_{i,j} &:= \int ([\widetilde{a}_{i,j}(r, x, y'', [X_r^{s_1 \wedge s_2,\xi,(m)}]) - \widetilde{a}_{i,j}(r, z, y'', [X_r^{s_1 \wedge s_2,\xi,(m)}])] \\
&\quad - [\widetilde{a}_{i,j}(r, x, y, [X_r^{s_1 \wedge s_2,\xi,(m)}]) - \widetilde{a}_{i,j}(r, z, y, [X_r^{s_1 \wedge s_2,\xi,(m)}])])\, \Delta_s \partial_x^{1+n} p_m(\mu, s, r, y, y'')\, dy'', \\
\mathrm{III}_{i,j} &:= \int\int \Delta_s [\widetilde{a}_{i,j}(r, x, y'', [X_r^{s,\xi,(m)}]) - \widetilde{a}_{i,j}(r, z, y'', [X_r^{s,\xi,(m)}])]\, \partial_y^n [\partial_\mu p_m(\mu, s_1 \vee s_2, r, x', y'')](y)\, dy''\, \mu(dx'), \\
\mathrm{IV}_{i,j} &:= \int\int [\widetilde{a}_{i,j}(r, x, y'', [X_r^{s_1 \wedge s_2,\xi,(m)}]) - \widetilde{a}_{i,j}(r, x, x', [X_r^{s_1 \wedge s_2,\xi,(m)}]) \\
&\quad - (\widetilde{a}_{i,j}(r, z, y'', [X_r^{s_1 \wedge s_2,\xi,(m)}]) - \widetilde{a}_{i,j}(r, z, x', [X_r^{s_1 \wedge s_2,\xi,(m)}]))].\Delta_s \partial_y^n [\partial_\mu p_m(\mu, s, r, x', y'')](y)\, dy''\, \mu(dx').
\end{aligned}$$

From (**HR**)(ii), (5.88), similarly to (5.95) with the map $\widetilde{a}_{i,j}$ instead of b_i, and (5.6), we get $|\mathrm{I}_{i,j}| \leq K|s_1 - s_2|^\beta (r - s_1 \vee s_2)^{-\frac{(1+n-\eta)}{2}-\beta}$ while, employing the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(r, x, y'', \mu)$, we obtain $|\mathrm{I}_{i,j}| \leq K|z-x|^\eta (r - s_1 \vee s_2)^{-\frac{(1+n)}{2}}$. From the two previous estimates and the space-time inequality (1.4), we thus obtain

$$\begin{aligned}
&\forall \beta \in [0,1],\ |\mathrm{I}_{i,j} H_2^{i,j} \left(\int_r^t a(v, z, [X_v^{s_1 \vee s_2,\xi,(m)}])\, dv, z - x \right) \widehat{p}_{m+1}(\mu, s_1 \vee s_2, r, t, x, z)| \\
&\qquad \leq K \left\{ \frac{1}{(t-r)^{1-\frac{\eta}{2}}(r - s_1 \vee s_2)^{\frac{1+n}{2}}} \wedge \frac{|s_1 - s_2|^\beta}{(t-r)(r - s_1 \vee s_2)^{\frac{1+n-\eta}{2}+\beta}} \right\} g(c(t-r), z-x) \qquad (5.98)
\end{aligned}$$

We now consider the two cases $|s_1 - s_2| \geq t - r$ and $|s_1 - s_2| \leq t - r$. In the first case, from the previous bound, we directly get

$$\begin{aligned}
&|\mathrm{I}_{i,j} H_2^{i,j} \left(\int_r^t a(v, z, [X_v^{s_1 \vee s_2,\xi,(m)}])\, dv, z - x \right) \widehat{p}_{m+1}(\mu, s_1 \vee s_2, r, t, x, z)| \\
&\qquad \leq \frac{|s_1 - s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}+\beta}(r - s_1 \vee s_2)^{\frac{1+n}{2}}} g(c(t-r), z-x)
\end{aligned}$$

while, in the second case, using (5.98) with $\beta = \frac{\eta}{2} \in [0,1]$ and the inequality $|s_1 - s_2| \leq t - r$, we obtain

$$\begin{aligned}
&|\mathrm{I}_{i,j} H_2^{i,j} \left(\int_r^t a(v, z, [X_v^{s_1 \vee s_2,\xi,(m)}])\, dv, z - x \right) \widehat{p}_{m+1}(\mu, s_1 \vee s_2, r, t, x, z)| \\
&\qquad \leq \frac{|s_1 - s_2|^\beta}{(t-r)^{1-\frac{\eta}{2}+\beta}(r - s_1 \vee s_2)^{\frac{1+n}{2}}} g(c(t-r), z-x).
\end{aligned}$$

We thus conclude

$$\begin{aligned}
&|\mathrm{I}_{i,j} H_2^{i,j} \left(\int_r^t a(v, z, [X_v^{s_1 \vee s_2,\xi,(m)}])\, dv, z - x \right) \widehat{p}_{m+1}(\mu, s_1 \vee s_2, r, t, x, z)| \\
&\quad \leq K \left\{ \frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r - s_1 \vee s_2)^{\frac{1+n}{2}}} \wedge \frac{1}{(t-r)(r - s_1 \vee s_2)^{\frac{1+n-\eta}{2}+\beta}} \right\} |s_1 - s_2|^\beta\, g(c(t-r), z-x)
\end{aligned}$$

where $\beta \in [0, \frac{\eta}{2})$. From (5.18) and the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(t,x,y,\mu)$, we get

$$|\mathrm{II}_{i,j}| \le K(|y-x|^\eta \wedge 1)\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta}}$$

while employing the η-Hölder regularity of $y \mapsto \widetilde{a}_{i,j}(t,x,y,\mu)$, we get

$$|\mathrm{II}_{i,j}| \le K\frac{|s_1-s_2|^\beta}{(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}$$

so that

$$\begin{aligned}|\mathrm{II}_{i,j} H_2^{i,j}&\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)|\\ &\le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta}}\wedge\frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x).\end{aligned}$$

We now distinguish the two cases $r \in [\frac{t+s_1\vee s_2}{2}, t)$ and $r \in [s_1\vee s_2, \frac{t+s_1\vee s_2}{2}]$ in the previous inequality. In the first case, we bound the minimum appearing in the right-hand side of the above inequality by the first argument and use the inequality $(r-s_1\vee s_2)^\beta \ge (t-r)^\beta$, while in the second case, we bound the minimum by the second argument and use the inequality $(r-s_1\vee s_2)^{\frac{\eta}{2}-\beta} \le (t-r)^{\frac{\eta}{2}-\beta}$. We thus obtain

$$\begin{aligned}|\mathrm{II}_{i,j} H_2^{i,j}&\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)|\\ &\le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r-s_1\vee s_2)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x).\end{aligned}$$

We deal with $\mathrm{III}_{i,j}$ similarly to $\mathrm{I}_{i,j}$ except that we use the estimate (5.12) instead of (5.6). Skipping technical details, we obtain

$$\begin{aligned}|\mathrm{III}_{i,j} H_2^{i,j}&\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)|\\ &\le K\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r-s_1\vee s_2)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x).\end{aligned}$$

For $\mathrm{IV}_{i,j}$, in the one hand, from the η-Hölder regularity of $x \mapsto \widetilde{a}_{i,j}(t,x,y,\mu)$, one gets $|\mathrm{IV}_{i,j}| \le K|s_1-s_2|^\beta|z-x|^\eta v_m^n(s_1,s_2,r)$ while, in the other hand, from the η-Hölder regularity of $y \mapsto \widetilde{a}_{i,j}(t,x,y,\mu)$, one gets $|\mathrm{IV}_{i,j}| \le K|s_1-s_2|^\beta u_m^n(s_1,s_2,r)$. Hence, from the space-time inequality (1.4), we conclude

$$\begin{aligned}|\mathrm{IV}_{i,j} H_2^{i,j}&\left(\int_r^t a(v,z,[X_v^{s_1\vee s_2,\xi,(m)}])\,dv, z-x\right)\widehat{p}_{m+1}(\mu,s_1\vee s_2,r,t,x,z)|\\ &\le K\left\{\frac{v_m^n(s_1,s_2,r)}{(t-r)^{1-\frac{\eta}{2}}}\wedge\frac{u_m^n(s_1,s_2,r)}{t-r}\right\}|s_1-s_2|^\beta\, g(c(t-r),z-x).\end{aligned}$$

Gathering the previous bound, we obtain

$$\begin{aligned}|\mathrm{III}_1| \le K&\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r-s_1\vee s_2)^{\frac{1+n}{2}}}\wedge\frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}+\frac{v_m^n(s_1,s_2,r)}{(t-r)^{1-\frac{\eta}{2}}}\wedge\frac{u_m^n(s_1,s_2,r)}{t-r}\right\}\\ &\times|s_1-s_2|^\beta\, g(c(t-r),z-x)\end{aligned}$$

for all $\beta \in [0, \frac{\eta}{2})$. In order to deal with III_2 and III_3, we employ (5.46) (note that we can bound $(r-s)^{\frac{1+n-\eta}{2}}u_m^n(s,r)$ and $(r-s)^{\frac{1+n}{2}}v_m^n(s,r)$ by K) to bound the quantity $\partial_y^n[\partial_\mu[a_{i,j}(r,x,[X_r^{s_1\wedge s_2,\xi,(m)}]) - a_{i,j}(r,z,[X_r^{s_1\wedge s_2,\xi,(m)}])]](y) := \mathrm{J}_{i,j}(y)$ as well as the mean value theorem and (5.95) with the map $a_{i,j}$ instead of b_i to bound $\Delta_s H_2^{i,j}\left(\int_r^t a(v,z,[X_v^{s,\xi,(m)}])\,dv, z-x\right)$ and $\Delta_s\widehat{p}_{m+1}(\mu,s,r,t,x,z)$. For both quantities, we obtain

$$\begin{aligned}|\mathrm{III}_3|+|\mathrm{III}_2| \le K&\left\{\frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}}}\wedge\frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\right\}\\ &\times|s_1-s_2|^\beta\, g(c(t-r),z-x)\end{aligned}$$

for all $\beta \in [0,1]$. Gathering the three previous estimates on III_1, III_2 and III_3 completes the proof of the announced estimate on III.

The induction hypothesis allows to conclude

$$
\begin{aligned}
&|\Delta_s \partial_v^n[\partial_\mu \mathcal{H}_{m+1}(\mu,s,r,t,x,z)](v)| \\
&\quad \le K\Bigg\{ \frac{1}{(t-r)^{1-\frac{\eta}{2}+\beta}(r-s_1\vee s_2)^{\frac{1+n}{2}}} \wedge \frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} \\
&\quad + \left(\frac{1}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} \wedge \frac{1}{(t-r)(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}} \right) \\
&\quad \times C_{m,n}(s_1\vee s_2, r) + \frac{1}{(t-r)^{2-\frac{\eta}{2}}} \int_r^t C_{m,n}(s_1\vee s_2, v)(v-s_1\vee s_2)^{-\frac{(1+n)}{2}-\beta+\eta}\, dv \Bigg\} \\
&\qquad \times |s_1-s_2|^\beta\, g(c(t-r), z-x).
\end{aligned} \tag{5.99}
$$

With the above estimates at hand we can now provide upper-bounds for the different terms appearing in the right-hand side of (5.89). The first estimate is given by (5.92). We thus consider the quantity $\Delta_s[\partial_v^n[\partial_\mu \widehat{p}_{m+1}] \otimes \Phi_{m+1}](\mu,s,t,x,z)$ and use the following decomposition

$$
\Delta_s[\partial_v^n \partial_\mu[\widehat{p}_{m+1}] \otimes \Phi_{m+1}](\mu,s,t,x,z)(v) = \mathrm{I} + \mathrm{II} + \mathrm{III}
$$

with

$$
\mathrm{I} := \int_{s_1\vee s_2}^t \int \Delta_s \partial_v^n[\partial_\mu \widehat{p}_{m+1}(\mu,s,r,x,y)](v) \Phi_{m+1}(\mu, s_1\vee s_2, r, t, y, z)\, dy\, dr,
$$

$$
\mathrm{II} := \int_{s_1\vee s_2}^t \int \partial_v^n[\partial_\mu \widehat{p}_{m+1}(\mu,s_1\wedge s_2,r,x,y)](v) \Delta_s \Phi_{m+1}(\mu,s,r,t,y,z)\, dy\, dr,
$$

and

$$
\mathrm{III} := -\int_{s_1\wedge s_2}^{s_1\vee s_2} \int \partial_v^n[\partial_\mu \widehat{p}_{m+1}(\mu,s_1\wedge s_2,r,x,y)](v) \Phi_{m+1}(\mu,s_1\wedge s_2,r,t,y,z)\, dy\, dr.
$$

From (5.92) and using the fact that $v \mapsto C_{m,n}(s_1\vee s_2, v)$ is non-decreasing we derive

$$
\begin{aligned}
|\mathrm{I}| &\le K \frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}} g(c(t-s_1\wedge s_2), z-x) \\
&\quad + K|s_1-s_2|^\beta \int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\, dr\, g(c(t-s_1\wedge s_2), z-x).
\end{aligned}
$$

From (5.97) and (5.41), one gets

$$
\forall \beta \in [0,1], \quad |\mathrm{II}| \le K \frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}} g(c(t-s_1\wedge s_2), z-x).
$$

Finally, using (5.41) together with the fact that $\beta \in [0, \frac{1+\eta}{2})$ if $n=0$ and $\beta \in [0, \eta/2)$ if $n=1$, one obtains

$$
|\mathrm{III}| \le K \frac{|s_1-s_2|^{\frac{1-n+\eta}{2}}}{(t-s_1\vee s_2)^{1-\frac{\eta}{2}}} g(c(t-s_1\wedge s_2), z-x) \le K \frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}} g(c(t-s_1\wedge s_2), z-x)
$$

where we used the inequality $|s_1-s_2| \le t - s_1\vee s_2$. Gathering the previous estimates finally yields

$$
\begin{aligned}
|\Delta_s[\partial_v^n \partial_\mu[\widehat{p}_{m+1}] \otimes \Phi_{m+1}](\mu,s,t,x,z)(v)| &\le K \frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}} g(c(t-s_1\wedge s_2), z-x) \\
&\quad + K|s_1-s_2|^\beta \int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\, dr\, g(c(t-s_1\wedge s_2), z-x).
\end{aligned} \tag{5.100}
$$

We now turn our attention to the term $\Delta_s(p_{m+1} \otimes \partial_v^n[\partial_\mu \mathcal{H}_{m+1}])(\mu,s,t,x,z)(v)$ and make use of a similar decomposition, namely

$$
\Delta_s(p_{m+1} \otimes \partial_v^n[\partial_\mu \mathcal{H}_{m+1}])(\mu,s,t,x,z)(v) = \mathrm{I} + \mathrm{II} + \mathrm{III}
$$

with

$$
\mathrm{I} := \int_{s_1\vee s_2}^t \int \Delta_s p_{m+1}(\mu,s,r,x,y) \partial_v^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s_1\vee s_2, r, t, y, z)](v)\, dy\, dr,
$$

$$\mathrm{II} := \int_{s_1\vee s_2}^t \int p_{m+1}(\mu, s_1\wedge s_2, r, x, y)\Delta_s\partial_v^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s, r, t, y, z)(v)\,dy\,dr,$$

and

$$\mathrm{III} := -\int_{s_1\wedge s_2}^{s_1\vee s_2} \int p_{m+1}(\mu, s_1\wedge s_2, r, x, y)\partial_v^n[\partial_\mu \mathcal{H}_{m+1}(\mu, s_1\wedge s_2, r, t, y, z)](v)\,dy\,dr.$$

From (5.47) (bounding $C_{m,n}$ by K) and (5.88), breaking the time integral into the two intervals $[s_1\vee s_2, (t+s_1\vee s_2)/2)$ and $[(t+s_1\vee s_2)/2, t]$ to balance the time singularity, after some standard computations, we obtain

$$\forall \beta\in[0,1],\ |\mathrm{I}| \le K\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_1), z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_2), z-x)\right\}.$$

To deal with II, we employ the estimate (5.99). For the first term appearing in the right-hand side of (5.99), we break the time integral into two intervals similarly to the previous estimate in order to balance the time singularity. For the second term, we bound the minimum of the two terms by the first one, namely $(t-r)^{-1+\frac{\eta}{2}}(r-s_1\vee s_2)^{-\frac{(1+n-\eta)}{2}-\beta}$ while for the third term we use Fubini's theorem. After some standard computations, we obtain

$$\begin{aligned}\forall \beta\in[0,\frac{\eta}{2}),\ |\mathrm{II}| \le K&\left\{\frac{1}{(t-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} + \int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\,dr\right\}\\ &\times |s_1-s_2|^\beta g(c(t-s_1\wedge s_2), z-x).\end{aligned}$$

Finally, using (5.47) (bounding $C_{m,n}$ by K) and breaking again the time integral into two intervals, we get

$$|\mathrm{III}| \le K\frac{|s_1-s_2|^{\frac{1-n+\eta}{2}}}{t-s_1\vee s_2} g(c(t-s_1\wedge s_2), z-x) \le K\frac{|s_1-s_2|^\beta}{(t-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_1\wedge s_2), z-x)$$

for all $\beta\in[0,\frac{\eta}{2})$. Gathering the three previous estimates finally yields

$$\begin{aligned}&|\Delta_s(p_{m+1}\otimes\partial_v^n[\partial_\mu\mathcal{H}_{m+1}])(\mu,s,t,x,z)(v)|\\ &\quad\le K\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_1), z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}} g(c(t-s_2), z-x)\right\}\\ &\quad + K|s_1-s_2|^\beta\int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\,dr\, g(c(t-s_1\wedge s_2), z-x).\end{aligned} \tag{5.101}$$

For the last term, namely $\Delta_s((p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}])\otimes\Phi_{m+1})(\mu,s,t,x,z)(v)$, as previously done, we decompose it as the sum of three terms I, II and III in a completely analogous way as the previous term. We make use of (5.101), (5.97) and (5.48). Skipping some technical details, we obtain

$$\begin{aligned}&|\Delta_s((p_{m+1}\otimes\partial_y^n[\partial_\mu\mathcal{H}_{m+1}])\otimes\Phi_{m+1})(\mu,s,t,x,z)(v)|\\ &\quad\le K\left\{\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n}{2}+\beta-\eta}} g(c(t-s_1), z-x) + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n}{2}+\beta-\eta}} g(c(t-s_2), z-x)\right\}\\ &\quad + K|s_1-s_2|^\beta\int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\,dr\, g(c(t-s_1\wedge s_2), z-x).\end{aligned} \tag{5.102}$$

Coming back to (5.89) and gathering the estimates (5.92), (5.100), (5.101) and (5.102) finally yield

$$\begin{aligned}&|\Delta_s\partial_y^n[\partial_\mu p_{m+1}(\mu,s,t,x,z)](v)|\\ &\le K\Big(\Big[\frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{1+n-\eta}{2}+\beta}} + \frac{|s_1-s_2|^\beta}{(t-s_1)^{\frac{\eta}{2}}}\int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\,dr\Big] g(c(t-s_1), z-x)\\ &\quad + \Big[\frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{1+n-\eta}{2}+\beta}} + \frac{|s_1-s_2|^\beta}{(t-s_2)^{\frac{\eta}{2}}}\int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2, r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\,dr\Big] g(c(t-s_2), z-x)\Big)\end{aligned}$$

so that

$$u^n_{m+1}(s_1,s_2,t) \le K\Big(\frac{1}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}} + \int_{s_1\vee s_2}^t \frac{C_{m,n}(s_1\vee s_2,r)}{(t-r)^{1-\frac{\eta}{2}}(r-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\,dr\Big)$$
$$\le \frac{K}{(t-s_1\vee s_2)^{\frac{1+n}{2}+\beta-\eta}}\left\{B\left(\frac{\eta}{2},\frac{1-n}{2}-\beta+\eta\right)\right.$$
$$\left.+\sum_{k=1}^m C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}-\beta+(i-1)\frac{\eta}{2}\right)(t-s_1\vee s_2)^{k\frac{\eta}{2}}\right\}$$

and similarly

$$v^n_{m+1}(s_1,s_2,t) \le \frac{K}{(t-s_1\vee s_2)^{\frac{1+n-\eta}{2}+\beta}}\left\{B\left(\frac{\eta}{2},\frac{1-n}{2}-\beta+\eta\right)\right.$$
$$\left.+\sum_{k=1}^m C^k\prod_{i=1}^{k+1}B\left(\frac{\eta}{2},\frac{1-n+\eta}{2}-\beta+(i-1)\frac{\eta}{2}\right)(t-s_1\vee s_2)^{k\frac{\eta}{2}}\right\}.$$

In a completely analogous manner as for the previous estimates, we thus derive that the induction hypothesis remains valid at step $m+1$. Coming back to (5.89) and bounding each term using our estimates combined with the two previous bounds, from the asymptotics of the Beta function, we deduce that (5.19) is valid at step $m+1$. The proof of the proposition is now complete.

6. Solving the related PDE on the Wasserstein space

This section is devoted to the proof of Theorem 3.10. Thanks to the regularity properties provided by Theorem 3.9, we are able to tackle the Cauchy problem (1.2) on any strip $[0,T]$. We first start with the following Proposition.

Proposition 6.1. Under the assumptions of Theorem 3.10, the mapping $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(t,x,\mu)\mapsto U(t,x,\mu)$ defined by (3.26) is continuous, belongs to $\mathcal{C}^{1,2,2}([0,T)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$, satisfies (3.27) and for any $(t,x,v,\mu)\in[0,T)\times(\mathbb{R}^d)^2\times\mathcal{P}_2(\mathbb{R}^d)$

$$\left|\partial_v^n[\partial_\mu U(t,x,\mu)](v)\right| \le C(T-t)^{-\frac{(1+n)}{2}}\exp(k|x|^2)(1+|v|^2+M_2^q(\mu)),\ n=0,1, \tag{6.1}$$

where $C:=C(T,|b|_\infty,[b]_H,|\widetilde{b}|_\infty,[\widetilde{b}]_H,|a|_\infty,[a]_H,|\widetilde{a}|_\infty,[\widetilde{a}]_H,\lambda,\eta)$, $k:=k(T,\lambda,\alpha)$ are positive constants. Moreover, U is a solution to the Cauchy problem (1.2) on the strip $[0,T]$.

Proof. We first remark that if $(\mu_n)_{n\ge1}$ is a sequence of $\mathcal{P}_2(\mathbb{R}^d)$ and if $(t_n)_{n\ge1}$ is a sequence of $[0,T)$ both satisfying $\lim_n|t_n-t|=\lim_n W_2(\mu_n,\mu)=0$, for some $(t,\mu)\in[0,T)\times\mathcal{P}_2(\mathbb{R}^d)$, then, by weak uniqueness, $([X_T^{t_n,\xi_n}])_{n\ge1}$ weakly converges to $[X_T^{t,\xi}]$, where $[\xi_n]=\mu_n$ and $[\xi]=\mu$, so that, passing to the limit in the parametrix infinite series (3.10) and using the relation (3.14), we deduce that $\int_{\mathbb{R}^d}|z|^2p(\mu_n,t_n,T,z)\,dz\to\int_{\mathbb{R}^d}|z|^2p(\mu,t,T,z)\,dz$ which in turn yields $\lim_n W_2([X_T^{t,\xi_n}],[X_T^{t,\xi}])=0$. We thus deduce that the two maps $[0,T]\times\mathcal{P}_2(\mathbb{R}^d)\ni(t,\mu)\mapsto h(z,[X_T^{t,\xi}])$, $[0,s]\times\mathcal{P}_2(\mathbb{R}^d)\ni(t,\mu)\mapsto f(s,z,[X_s^{t,\xi}])$ are continuous so that the mapping $[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(t,x,\mu)\mapsto U(t,x,\mu)$ is also continuous.

We now prove that $\mathcal{P}_2(\mathbb{R}^d)\ni(x,\mu)\mapsto U(t,x,\mu)\in\mathcal{C}^{2,2}(\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d))$, for $t\in[0,T)$ and that $[0,T)\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\ni(t,x,\mu)\mapsto\mathcal{L}_tU(t,x,\mu)$ is continuous, where the operator $\mathcal{L}_t$ is defined by (1.3).

From Theorem 3.9 and the relation (3.14), the map $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto p(\mu,t,T,z)$ is partially $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ (see Chapter 5 of [CD18] for a definition of partial $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ regularity) with derivatives given by

$$\partial_v^n[\partial_\mu p(\mu,t,T,z)](v)=\partial_x^{1+n}p(\mu,t,T,v,z)+\int_{\mathbb{R}^d}\partial_v^n[\partial_\mu p(\mu,t,T,x,z)](v)\,\mu(dx),\quad n=0,\,1.$$

From assumption (**HST**), we thus deduce that the two maps $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto h(z,[X_T^{t,\xi}])$, $\mathcal{P}_2(\mathbb{R}^d)\ni\mu\mapsto f(s,z,[X_s^{t,\xi}])$ are partially $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ for any fixed T, $s>t$ and $z\in\mathbb{R}^d$. Moreover, by Fubini's theorem, their Lions derivatives are given by

$$\partial_v^n[\partial_\mu h(z,[X_T^{t,\xi}])](v)=\int_{\mathbb{R}^d}\widetilde{h}(z,y,[X_T^{t,\xi}])\,\partial_x^{1+n}p(\mu,t,T,v,y)\,dy$$
$$+\int_{(\mathbb{R}^d)^2}\widetilde{h}(z,y,[X_T^{t,\xi}])\,\partial_v^n[\partial_\mu p(\mu,t,T,x,y)](v)\,dy\,\mu(dx) \tag{6.2}$$

and by Lemma (2.1)

$$\partial_v^n[\partial_\mu f(s,z,[X_s^{t,\xi}])](v) = \int_{\mathbb{R}^d} [\widetilde{f}(s,z,y,[X_s^{t,\xi}]) - \widetilde{f}(s,z,v,[X_s^{t,\xi}])]\, \partial_x^{1+n} p(\mu,t,s,v,y)\, dy$$
$$+ \int_{(\mathbb{R}^d)^2} \widetilde{f}(s,z,y,[X_s^{t,\xi}])\, \partial_v^n[\partial_\mu p(\mu,t,s,x,y)](v)\, dy\, \mu(dx). \tag{6.3}$$

We may break the integral appearing in the right-hand side of (6.3) into two parts J_1 and J_2 by dividing the domain of integration into two domains. In the first part J_1, the dy-integration is taken over a bounded domain D containing v such that $|y-v| \geq 1$ if $y \notin D$. Using the η-Hölder regularity of $y \mapsto \widetilde{f}(s,z,y,[X_s^{t,\xi}])$ on D, (3.12) and the space-time inequality (1.4), we get

$$|\mathrm{J}_1| \leq C(s-t)^{\frac{-1-n+\eta}{2}}.$$

As for J_2, for $\alpha < c := c(\lambda)$, where c is the constant appearing in (3.12), from (3.25), the space-time inequality (1.4) and noting that $M_2([X_T^{t,\xi}]) \leq C(1+M_2(\mu))$, we obtain

$$|\mathrm{J}_2| \leq C \exp\left(\alpha \frac{|z|^2}{T}\right) \int_{|y-v|\geq 1} (s-t)^{-1}(1+|y|^2+M_2^q(\mu))\, g(c(s-t), y-v)\, dy$$
$$\leq C \exp\left(\alpha \frac{|z|^2}{T}\right) (s-t)^{\frac{-1-n+\eta}{2}} (1+|v|^2+M_2^q(\mu)).$$

Also, from (3.15) and (3.25), we derive

$$\Big| \int_{(\mathbb{R}^d)^2} \widetilde{f}(s,z,y,[X_s^{t,\xi}])\, \partial_v^n[\partial_\mu p(\mu,t,s,x,y)](v)\, dy\, \mu(dx) \Big| \leq C \exp\left(\alpha \frac{|z|^2}{T}\right) (s-t)^{\frac{-1-n+\eta}{2}} (1+M_2^q(\mu)).$$

Gathering the previous estimates, we obtain

$$|\partial_v^n[\partial_\mu f(s,z,[X_s^{t,\xi}])](v)| \leq C \exp\left(\alpha \frac{|z|^2}{T}\right) (s-t)^{\frac{-1-n+\eta}{2}} (1+|v|^2+M_2^q(\mu)). \tag{6.4}$$

From (6.2), (3.12), (3.25) and similar computations

$$|\partial_v^n[\partial_\mu h(z,[X_T^{t,\xi}])](v)| \leq C \exp\left(\alpha \frac{|z|^2}{T}\right) (T-t)^{-\frac{(1+n)}{2}} (1+|v|^2+M_2^q(\mu)). \tag{6.5}$$

The estimates (3.12), (3.15) and (6.4) allow to conclude that the map $(x,\mu) \mapsto U(t,x,\mu)$ is in $\mathcal{C}^{2,2}(\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$ with derivatives given by

$$\partial_v^n[\partial_\mu U(t,x,\mu)](v) = \int_{\mathbb{R}^d} h(z,[X_T^{t,\xi}])\, \partial_v^n[\partial_\mu p(\mu,t,T,x,z)](v)\, dz + \int_{\mathbb{R}^d} \partial_v^n[\partial_\mu h(z,[X_T^{t,\xi}])](v)\, p(\mu,t,T,x,z)\, dz$$
$$- \int_t^T \int_{\mathbb{R}^d} \partial_v^n[\partial_\mu f(s,z,[X_s^{t,\xi}])](v)\, p(\mu,t,s,x,z)\, dz\, ds \tag{6.6}$$
$$- \int_t^T \int_{\mathbb{R}^d} f(s,z,[X_s^{t,\xi}])\, \partial_v^n[\partial_\mu p(\mu,t,s,x,z)](v)\, dz\, ds$$

for $n = 0$, 1 and

$$\partial_x^n U(t,x,\mu) = \int_{\mathbb{R}^d} h(z,[X_T^{t,\xi}])\, \partial_x^n p(\mu,t,T,x,z)\, dz$$
$$- \int_t^T \int_{\mathbb{R}^d} [f(s,z,[X_s^{t,\xi}]) - f(s,x,[X_s^{t,\xi}])]\, \partial_x^n p(\mu,t,s,x,z)\, dz\, ds \tag{6.7}$$

for $n = 0$, 1, 2. Note that we may break the last integral appearing in the right-hand side of (6.7) into two parts by dividing the domain of integration into two domains as we did before. Then, using the local Hölder continuity of $z \mapsto f(s,z,\mu)$, (3.24), the estimate (3.12), we get

$$\Big| \int_{\mathbb{R}^d} [f(s,z,[X_s^{t,\xi}]) - f(s,x,[X_s^{t,\xi}])]\, \partial_x^{1+n} p(\mu,t,s,x,z)\, dz \Big|$$
$$\leq C(s-t)^{\frac{-1-n+\eta}{2}} \left\{ \int_{\mathbb{R}^d} e^{\alpha \frac{|z|^2}{T}} g(c(s-t), z-x)\, dz + e^{\alpha \frac{|x|^2}{T}} \right\} (1+M_2^q(\mu))$$
$$\leq C(s-t)^{\frac{-1-n+\eta}{2}} e^{k \frac{|x|^2}{T}} (1+M_2^q(\mu))$$

where we used the fact that the constant α is sufficiently small, namely $\alpha < c$, c being the constant appearing in (3.12) and the inequality: for any $\alpha > 0$, $\varepsilon > 0$, there exists a positive constant $C := C(\alpha, \varepsilon)$ such that for any $(z, x) \in (\mathbb{R}^d)^2$,

$$\alpha|z|^2 - (\alpha + \varepsilon)|z - x|^2 \leq C|x|^2. \tag{6.8}$$

The previous estimate as well as (6.4) and (3.15) ensure that the integrals appearing in (6.6) and (6.7) are well defined if α is sufficiently small. We thus conclude from (6.6) and (6.7) that $[0, T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \ni (t, x, \mu) \mapsto \mathcal{L}_t U(t, x, \mu)$ is continuous.

Finally, from (3.11) and (6.8), we get

$$\begin{aligned}|U(t,x,\mu)| &\leq C\left\{\int_{\mathbb{R}^d} \exp\Big(\alpha\frac{|z|^2}{T}\Big) g(c(T-t), z-x)\, dz + \int_t^T \int_{\mathbb{R}^d} \exp\Big(\alpha\frac{|z|^2}{T}\Big) g(c(s-t), z-x)\, dz\, ds\right\} \\ &\quad \times (1 + M_2^q(\mu)) \\ &\leq C \exp\Big(k|x|^2\Big)(1 + M_2^q(\mu)).\end{aligned}$$

The proof of (6.1) follows from (6.4), (6.5) and (6.6).

Let us now prove that U is in $\mathcal{C}^{1,2,2}([0, T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d))$. From the Markov property satisfied by the SDE (1.1) (which is inherited from the well-posedness of the associated martingale problem) we obtain the following identity for all $h > 0$

$$U(t - h, x, \mu) = \mathbb{E}\left[U(t, X_t^{t-h,x,\mu}, [X_t^{t-h,\xi}]) - \int_{t-h}^t f(r, X_r^{t-h,x,\mu}, [X_r^{t-h,\xi}])dr\right].$$

From (3.12), we clearly get $|\partial_x U(r, x, \mu)| \leq C(T - r)^{-1/2}$ so that, combining Proposition 6.1, especially the estimate (6.1), with the chain rule formula of Proposition 2.1 (with respect to the space and measure variables only) we get

$$\mathbb{E}\left[U(t, X_t^{t-h,x,\mu}, [X_t^{t-h,\xi}])\right] = U(t, x, \mu) + \mathbb{E}\left[\int_{t-h}^t \mathcal{L}_r U(t, X_r^{t-h,x,\mu}, [X_r^{t-h,\xi}])dr\right].$$

Hence

$$\frac{1}{h}\left(U(t - h, x, \mu) - U(t, x, \mu)\right) = \frac{1}{h}\mathbb{E}\left[\int_{t-h}^t \left\{\mathcal{L}_r U(t, X_r^{t-h,x,\mu}, [X_r^{t-h,\xi}]) - f(r, X_r^{t-h,x,\mu}, [X_r^{t-h,\xi}])\right\} dr\right]$$

and letting $h \downarrow 0$, from the boundedness and continuity of the coefficients, we deduce that U is left differentiable in time. Still from the continuity of the coefficients and f, we then conclude that it is differentiable in time with

$$\partial_t U(t, x, \mu) = -\mathcal{L}_t U(t, x, \mu) + f(t, x, \mu).$$

Hence, the map U solves the PDE (3.26). □

In order to get the uniqueness result, first fix any $0 \leq t \leq s < T$ and consider any solution V to the Cauchy problem (1.2) satisfying (2.4) on any interval $[0, T']$, with $T' < T$, as well as (3.27). We apply the chain rule formula of Proposition 2.1 to $\left\{V(s, X_s^{t,x,\mu}, [X_s^{t,\xi}]),\ t \leq s < T\right\}$ and use the fact that $(\partial_t + \mathcal{L}_t)V(t, x, \mu) = f(t, x, \mu)$, for $(t, x, \mu) \in [0, T) \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ to get

$$\begin{aligned}V(s, X_s^{t,x,\mu}, [X_s^{t,\xi}]) &= V(t, x, \mu) + \int_t^s f(r, X_r^{t,x,\mu}, [X_r^{t,\xi}])\, dr \\ &\quad + \int_t^s \sum_{i=1}^d \sum_{j=1}^q \sigma_{i,j}(r, X_r^{t,x,\mu}, [X_r^{t,\xi}])\, \partial_{x_i} V(r, X_r^{t,x,\mu}, [X_r^{t,\xi}])dB_r^j.\end{aligned}$$

The local martingale appearing in the right-hand side of the above equality is in fact a true martingale since $V(s, X_s^{t,x,\mu}, [X_s^{t,\xi}])$ and $\int_t^s f(r, X_r^{t,x,\mu}, [X_r^{t,\xi}])\, dr$ are both square integrable if the constant α and k appearing in the two conditions (3.24) and (3.27) are small enough, that is, α and k strictly less than $c/2$, $c := c(\lambda)$ being the constant appearing in (3.11) is sufficient.

Hence, taking expectation in the previous equality, then passing to the limit as $s \uparrow T$ and using the continuity assumption at the boundary, we obtain

$$V(t, x, \mu) = \mathbb{E}[h(X_T^{t,x,\mu}, [X_T^{t,\xi}]) - \int_t^T f(r, X_r^{t,x,\mu}, [X_r^{t,\xi}])\, dr]$$

which completes the proof of Theorem 3.10.

7. Appendix

7.1. **Weak existence for the SDE** (1.1)**.** We here provide a simple proof based on a compactness argument which allows to establish the existence of weak solutions to the SDE (1.1) under assumption (**HR**). We consider the sequence of probability measures $(\mathbb{P}^{(m)})_{m\geq 0}$ on $\mathcal{C}([0,\infty),\mathbb{R}^d)$ introduced in the beginning of Section 5.1, namely for any $\mathbb{P}^{(0)}\in\mathcal{C}([0,\infty),\mathcal{P}(\mathbb{R}^d))$, we let $\mathbb{P}^{(m+1)}$ be a probability measure induced by a weak solution to the SDE with dynamics

$$X_t^{\xi,(m+1)}=\xi+\int_0^t b(r,X_r^{\xi,(m+1)},\mathbb{P}^{(m)}(r))dr+\int_0^t\sigma(r,X_r^{\xi,(m+1)},\mathbb{P}^{(m)}(r))dW_r. \tag{7.1}$$

Note that the above SDE admits a weak solution since the two maps $\mathbb{R}_+\times\mathbb{R}^d\ni(t,x)\mapsto\widehat{b}(t,x):=b(t,x,\mathbb{P}^{(m)}(t))$, $\widehat{\sigma}(t,x):=\sigma(t,x,\mathbb{P}^{(m)}(t))$ are bounded and continuous under (**HC**). Since ξ is square integrable and the coefficients are bounded the sequence $(\mathbb{P}^{(m)})_{m\geq 0}$ is tight. Relabelling the indices if necessary, we may assert that $(\mathbb{P}^{(m)})_{m\geq 0}$ converges weakly to a probability measure $\mathbb{P}^\infty$. Our aim is to prove that $\mathbb{P}^\infty$ is a solution to the martingale problem of definition 3.1. For every continuous and bounded function f defined on $\mathbb{R}^d$, the weak convergence of $(\mathbb{P}^{(m)})_{m\geq 1}$ to $\mathbb{P}^\infty$ gives

$$\mathbb{E}_{\mathbb{P}^\infty}[f(y(0))]=\lim_{m\to\infty}\mathbb{E}_{\mathbb{P}^{(m)}}[f(y(0))]=\int f(y)\mu(dy)$$

so that $\mathbb{P}^\infty(y(0)\in\Gamma)=\mu(\Gamma)$, $\Gamma\in\mathcal{B}(\mathbb{R}^d)$. It remains to prove that

$$\mathbb{E}_{\mathbb{P}^\infty}\Big[\Big(f(y(t))-f(y(s))\\ -\int_s^t(\sum_{i,j}\frac{1}{2}a_{i,j}(r,y(r),\mathbb{P}^\infty(r))\partial_{x_i,x_j}f(y(r))+b_i(r,y(r),\mathbb{P}^\infty(r))\partial_{x_i}f(y(r))dr\Big)\mathscr{G}(y)\Big]=0$$

for any bounded $\mathcal{F}_s$-measurable function $\mathscr{G}:\mathcal{C}([0,\infty),\mathbb{R}^d)\to\mathbb{R}$. For every $m\geq 0$, the fact that $\mathbb{P}^{(m+1)}$ is the probability measure induced by a weak solution to the SDE (7.1) gives

$$\mathbb{E}_{\mathbb{P}^{(m+1)}}\Big[\Big(f(y(t))-f(y(s))\\ -\int_s^t(\sum_{i,j}\frac{1}{2}a_{i,j}(r,y(r),\mathbb{P}^{(m)}(r))\partial_{x_i,x_j}f(y(r))+b_i(r,y(r),\mathbb{P}^{(m)}(r))\partial_{x_i}f(y(r))dr\Big)\mathscr{G}(y)\Big]=0.$$

By weak convergence of $(\mathbb{P}^{(m)})_{m\geq 1}$, one gets

$$\mathbb{E}_{\mathbb{P}^\infty}\Big[(f(y(t))-f(y(s)))\mathscr{G}(y)]=\lim_{m\to\infty}\mathbb{E}_{\mathbb{P}^{(m)}}[(f(y(t))-f(y(s)))\mathscr{G}(y)]$$

and

$$\mathbb{E}_{\mathbb{P}^\infty}[\int_s^t(\sum_{i,j}\frac{1}{2}a_{i,j}(r,y(r),\mathbb{P}^\infty(r))\partial_{x_i,x_j}f(y(r))+b_i(r,y(r),\mathbb{P}^\infty(r))\partial_{x_i}f(y(r))dr)\mathscr{G}(y)]$$

$$=\lim_{m\to\infty}\mathbb{E}_{\mathbb{P}^{(m+1)}}\Big[\int_s^t(\sum_{i,j}\frac{1}{2}a_{i,j}(r,y(r),\mathbb{P}^\infty(r))\partial_{x_i,x_j}f(y(r))+b_i(r,y(r),\mathbb{P}^\infty(r))\partial_{x_i}f(y(r))dr\Big)\mathscr{G}(y)\ \Big]$$

so that, since $\mathscr{G}$ is bounded, it suffices to prove

$$\lim_{m\to\infty}\mathbb{E}_{\mathbb{P}^{(m+1)}}\Big[\int_s^t\Big|\sum_{i,j}\frac{1}{2}(a_{i,j}(r,y(r),\mathbb{P}^\infty(r))-a_{i,j}(r,y(r),\mathbb{P}^{(m)}(r)))\partial_{x_i,x_j}f(y(r))\\ +(b_i(r,y(r),\mathbb{P}^\infty(r))-b_i(r,y(r),\mathbb{P}^{(m)}(r)))\partial_{x_i}f(y(r))\Big|\,dr\ \Big]=0. \tag{7.2}$$

By weak convergence of $(\mathbb{P}^{(m)})_{m\geq 0}$,

$$\limsup_{m\to\infty}\mathbb{P}^{(m)}(y:|y(r)|\geq R)\leq\mathbb{P}^\infty(y:|y(r)|\geq R)$$

and choosing R large enough the right-hand side of the previous inequality is smaller than $\varepsilon/2$. Moreover, under (**HC**), again by weak convergence, for any $s'\in[s,t]$, one has

$$\sup_{|x|\leq R,s\leq r\leq t}|a_{i,j}(r,x,\mathbb{P}^\infty(s'))-a_{i,j}(r,x,\mathbb{P}^{(m)}(s'))|+\sup_{|x|\leq R,s\leq r\leq t}|b_i(r,x,\mathbb{P}^\infty(s'))-b_i(r,x,\mathbb{P}^{(m)}(s'))|\leq\frac{\varepsilon}{2}$$

for m large enough. We thus conclude that (7.2) is valid. This completes the proof.

7.2. **A technical lemma.**

Lemma 7.1. For all $h \in \mathcal{C}_b(\mathbb{R}^d)$ and all $(t,x,\mu) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, one has

$$\lim_{r \downarrow 0} \int_{\mathbb{R}^d} h(z)\, \bar{p}^z(\mu,t,t+r,x,z)\, dz = h(x).$$

Proof. We remind the readers that $\bar{p}^z(\mu,t,t+r,x,z) = g\left(\int_t^{t+r} a(s,z,[\bar{X}_s^{t,\xi,z}])ds, z-x\right)$, where ξ has law μ. It is important to note that in the previous expression the law $[\bar{X}_s^{t,\xi,z}]$ also depends on the terminal point z at which the diffusion coefficient is frozen. We introduce the density function $z \mapsto \bar{p}(\mu,t,s,z) = \int g\left(\int_s^t a(r,x,[\bar{X}_r^{t,\xi,x}])\, dr, z-x'\right) \mu(dx')$ of the random variable $\bar{X}_s^{t,\xi,x}$ given by $\bar{X}_s^{t,\xi,x} = \xi + \int_t^s \sigma(r,x,[\bar{X}_r^{t,\xi,x}])\, dW_r$. On the one hand, thanks to (**HR**)(ii) and Fubini's theorem, one gets

$$\begin{aligned}
\max_{i,j} |a_{i,j}(s,z,[\bar{X}_s^{t,\xi,z}]) - a_{i,j}(s,z,[\bar{X}_s^{t,\xi,x}])| & \\
&\leq \Big| \int A(s,z,y',[\bar{X}_s^{t,\xi,z}],[\bar{X}_s^{t,\xi,x}])(\bar{p}^z(\mu,t,s,y') - \bar{p}^x(\mu,t,s,y'))\, dy' \Big| \\
&\leq \Big| \int_{(\mathbb{R}^d)^2} (A(s,z,y',[\bar{X}_s^{t,\xi,z}],[\bar{X}_s^{t,\xi,x}]) - A(s,z,x',[\bar{X}_s^{t,\xi,z}],[\bar{X}_s^{t,\xi,x}])) \\
&\quad \times (\bar{p}^z(\mu,t,s,x',y') - \bar{p}^x(\mu,t,s,x',y'))\, dy'\, \mu(dx') \Big| \\
&\leq C(s-t)^{\eta/2}
\end{aligned}$$

where we used the fact that $y \mapsto A(s,z,y,\mu,\nu)$ is η-Hölder uniformly with respect to the other variables as well as the space-time inequality (1.4). On the other hand, again from (**HR**), one derives $|a(s,z,[\bar{X}_s^{t,\xi,x}]) - a(s,x,[\bar{X}_s^{t,\xi,x}])| \leq C(|z-x|^\eta \wedge 1)$. Hence, taking advantage of the two previous estimates, from the mean value theorem and the space-time inequality (1.4), we obtain:

$$\begin{aligned}
|\bar{p}^z(\mu,t,t+r,x,z) - \bar{p}^x(\mu,t,t+r,x,z)| & \\
&\leq \Big| g\Big(\int_t^{t+r} a(s,z,[\bar{X}_s^{t,\xi,z}])ds, z-x\Big) - g\Big(\int_t^{t+r} a(s,z,[\bar{X}_s^{t,\xi,x}])ds, z-x\Big) \Big| \\
&\quad + \Big| g\Big(\int_t^{t+r} a(s,z,[\bar{X}_s^{t,\xi,x}])ds, z-x\Big) - g\Big(\int_t^{t+r} a(s,x,[\bar{X}_s^{t,\xi,x}])ds, z-x\Big) \Big| \\
&\leq \frac{C}{r} \int_t^{t+r} \Big\{ (s-t)^{\frac{\eta}{2}} + (|z-x|^\eta \wedge 1) \Big\}\, ds\, g(cr, z-x) \\
&\leq C r^{\frac{\eta}{2}} g(cr, z-x).
\end{aligned}$$

Finally, from the previous inequality, we easily conclude

$$\int_{\mathbb{R}^d} h(z)\, \bar{p}^z(\mu,t,t+r,x,z)\, dz = \int_{\mathbb{R}^d} h(z)\, \bar{p}^x(\mu,t,t+r,x,z)\, dz + \mathcal{O}(r^{\frac{\eta}{2}}).$$

Passing to the limit as $r \downarrow 0$ in the previous equality concludes the proof. □

Acknowledgments.

For the first Author, this work has been partially supported by the ANR project ANR-15-IDEX-02.

References

[Bañ18] D. Baños. The Bismut–Elworthy–Li formula for mean-field stochastic differential equations. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 54(1):220–233, 02 2018.

[BFY17] A. Bensoussan, J. Frehse, and S.C.P. Yam. On the interpretation of the master equation. *Stochastic Processes and their Applications*, 127(7):2093 – 2137, 2017.

[BLPR17] R. Buckdahn, J. Li, S. Peng, and C. Rainer. Mean-field stochastic differential equations and associated pdes. *Ann. Probab.*, 45(2):824–878, 03 2017.

[BP03] R. F. Bass and E. A. Perkins. Degenerate stochastic differential equations with hölder continuous coefficients and super-markov chains. *Trans. Amer. Math. Soc.*, 355(1):373–405, 2003.

[BP09] R. F. Bass and E. A. Perkins. A new technique for proving uniqueness for martingale problems. *From Probability to Geometry (I): Volume in Honor of the 60th Birthday of Jean-Michel Bismut*, pages 47–53, 2009.

[Car13] P. Cardaliaguet. Notes on mean field games. `https://www.ceremade.dauphine.fr/cardaliaguet/MFG20130420.pdf`, 2013.

[CCD14] J.-F. Chassagneux, D. Crisan, and F. Delarue. A probabilistic approach to classical solutions of the master equation for large population equilibria. ArXiv e-prints arXiv:1411.3009, 2014.

[CD18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games*. Probability Theory and Stochastic Modelling. Springer International Publishing, 2018.

[CdR15] P.-E. Chaudru de Raynal. Strong well-posedness of mckean-vlasov stochastic differential equations with hölder drift. ArXiv e-prints arXiv:1512.08096v2, 2015.

[CM17] D. Crisan and E. McMurray. Smoothing properties of mckean–vlasov sdes. *Probability Theory and Related Fields*, Apr 2017.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[FL17] N. Frikha and L. Li. Weak uniqueness and density estimates for sdes with coefficients depending on some path-functionals. *Preprint*, 2017.

[Fri64] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall, 1964.

[Fri11] A. Friedman. *Stochastic Differential Equations and Applications*, pages 75–148. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[Fri17] N. Frikha. Stochastic approximation, markovian perturbation of stochastic processes and their applications. Hdr report, Université Paris Diderot, 2017.

[Fun84] T. Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(3):331–348, Oct 1984.

[Gär88] J. Gärtner. On the Mckean-Vlasov Limit for Interacting Diffusions. *Mathematische Nachrichten*, 137(1):197–248, 1988.

[HHJ15] M. Hairer, M. Hutzenthaler, and A. Jentzen. Loss of regularity for Kolmogorov equations. *Ann. Probab.*, 43(2):468–527, 03 2015.

[Hör67] L. Hörmander. Hypoelliptic second order differential operators. *Acta. Math.*, 119:147–171, 1967.

[HvS18] W. Hammersley, D. Šiška, and L. Szpruch. McKean-Vlasov SDEs under Measure Dependent Lyapunov Conditions. Preprint Arxiv arXiv:1802.03974, 2018.

[Jou97] B. Jourdain. Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized burgers' equations. *ESAIM: PS*, 1:339–355, 1997.

[Kac56] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Contributions to Astronomy and Physics*, pages 171–197, Berkeley, Calif., 1956. University of California Press.

[Kes60] H. Kestelman. *Modern theories of integration*. 2nd revised ed. Dover Publications, Inc., New York, 1960.

[KM00] V. Konakov and E. Mammen. Local limit theorems for transition densities of Markov chains converging to diffusions. *Prob. Th. Rel. Fields*, 117:551–587, 2000.

[Kry99] N. V. Krylov. *On Kolmogorov's equations for finite dimensional diffusions*, pages 1–63. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[Lac18] D. Lacker. On a strong form of propagation of chaos for mckean-vlasov equations. Preprint Arxiv arXiv:1805.04476, 2018.

[Lio14] P.-L. Lions. Cours au collège de France. `http://www.college-de-france.fr/site/pierre-louis-lions/seminar-2014-11-14-11h1` 2014.

[LM16] J. Li and H. Min. Weak solutions of mean-field stochastic differential equations and application to zero-sum stochastic differential games. *SIAM Journal on Control and Optimization*, 54(3):1826–1858, 2016.

[McK66] H. P. McKean. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6):1907–1911, 12 1966.

[McK67] H. P. McKean. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57, 1967.

[Men11] S. Menozzi. Parametrix techniques and martingale problems for some degenerate kolmogorov equations. *Electron. Commun. Probab.*, 16:234–250, 2011.

[MS67] H. P. McKean and I. M. Singer. Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry*, 1:43–69, 1967.

[MV18] Y. S. Mishura and A. Y. Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. Preprint Arxiv arXiv:1603.02212, 2018.

[Oel84] K. Oelschlager. A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.*, 12(2):458–479, 05 1984.

[SV79] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, 1979.

[Szn91] A.-S. Sznitman. Topics in propagation of chaos. In Paul-Louis Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*, pages 165–251, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.

[Tan78] H. Tanaka. Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 46(1):67–105, 1978.

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

Paul-Éric Chaudru de Raynal, Université de Savoie Mont Blanc, CNRS, LAMA, F-73000 Chambéry, France
E-mail address: pe.deraynal@univ-smb.fr

Noufel Frikha, Laboratoire de Probabilités, Statistique et Modélisation (LPSM), Sorbonne Paris Cité, UMR CNRS 8001, Université Paris Diderot, F-75205 Paris, France
E-mail address: frikha@math.univ-paris-diderot.fr