RATE OF CONVERGENCE OF A PARTICLE METHOD TO THE SOLUTION OF THE MC KEAN - VLASOV'S EQUATION

FABIO ANTONELLI

Università della Tuscia
Dipartimento di Scienze Ambientali,
Via S. Camillo De Lellis
01100 Viterbo - ITALY.

ARTURO KOHATSU-HIGA

Universitat Pompeu Fabra
Departament d'Economia
Ramon Trias Fargas 25-27,
Barcelona 08005 - SPAIN

ABSTRACT. This paper studies the rate of convergence of an appropriate discretization scheme of the solution of the Mc Kean - Vlasov equation introduced by Bossy and Talay. More specifically, we consider approximations of the distribution and of the density of the solution of the stochastic differential equation associated to the Mc Kean - Vlasov equation. The scheme adopted here is a mixed one: Euler/weakly interacting particle system. If n is the number of weakly interacting particles and h is the uniform step in the time discretization, we prove that the rate of convergence of the distribution functions of the approximating sequence in the $L^1(\Omega \times \mathbb{R})$ norm is of the order of $\frac{1}{\sqrt{n}} + h$, while for the densities is of the order $h + \frac{1}{\sqrt{n\,h}}$. The rates of convergence with respect to the supremum norm are also calculated. This result is obtained by carefully employing techniques of Malliavin Calculus.

1. Introduction

In a series of articles (see [BT1] , [BT2], [Ta]), Bossy and Talay study the numerical approximation of the solutions to the McKean - Vlasov equation and to the Burgers equation. The McKean - Vlasov equation is obtained as the diffusive limit of a particle system, describing the behaviour of a high density gas. Its solution is a probability law/density and it can be represented as the law of the solution of an associated nonlinear stochastic differential equation (for further details we refer the reader to [G]).

In their paper, Bossy and Talay choose to approximate the McKean - Vlasov limit by replicating the behaviour with a system of n weakly interacting particles, each following a sde discretized in time with step $h \in (0,1]$. In [BT1] it is proved that when $n \to \infty$ and $h \to 0$, then the empirical distribution function of these n particles converges, in an weak sense to be defined later, towards the solution of the McKean - Vlasov limit with a rate at

1991 *Mathematics Subject Classification.* Primary: 60H10, 60K35 Secondary: 34F05 90A12.
Key words and phrases. Mc Kean - Vlasov equation, Malliavin Calculus.

least of the order $\frac{1}{\sqrt{n}} + \sqrt{h}$. Through some simulations it can be clearly seen that the rate in n is optimal but that the rate in h is probably better than $\sqrt{h}$.

In this article, we prove that the rate of convergence of the scheme constructed by Bossy and Talay is actually at least of the order $\frac{1}{\sqrt{n}} + h$, as they also suspected on the basis of some numerical simulations they ran.

To make our introduction more precise, we recall that the McKean - Vlasov equation can be described by means of four Lipschitz kernels $a(x,y)$, $b(x,y)$, $f(x,y)$ and $g(x,y)$ from $\mathbb{R}^2$ to $\mathbb{R}$ and of a differential operator, acting on the probability measures, defined by

$$L(\mu)h(x) = \frac{1}{2}\left[b(x, \int_{\mathbb{R}} g(x,y)d\mu(y))\right]^2 h''(x) + \left[a(x, \int_{\mathbb{R}} f(x,y)d\mu(y))\right] h'(x).$$

A family of probability measure $\{\mu_t\}_{t\geq 0}$ is said to be the solution of the McKean - Vlasov equation if it solves

$$(1.1) \qquad \begin{cases} \dfrac{d}{dt} < \mu_t, h >=< \mu_t, L(\mu_t)h >, \quad \forall\, h \in C_K^\infty(\mathbb{R}), \text{ (compact support)} \\ \mu_{t=0} = \mu_0, \end{cases}$$

where μ_0 is an initial probability measure. Applications and a general discussion about the above equation can be found in Gärtner ([G]). For example, this kind of system is the natural limit equation for the theory of propagation of chaos that allows the interpretation of the above system as the equation satisfied by the limit of the empirical measure associated with a random particle system with long-range weak interaction. Applications of this model arise in Newtonian physics, statistical mechanics, chemical kinetics, segregation problems in biological populations between others .

By associating a martingale problem to the operator L, μ_t can also be characterized through the stochastic differential equation (sde)

$$(1.2) \qquad X_t = \xi + \int_0^t a(X_s, \int_{\mathbb{R}} f(X_s,y)d\mu_s(y))ds + \int_0^t b(X_s, \int_{\mathbb{R}} g(X_s,y)d\mu_s(y)) \cdot dW_s,$$

where μ_t denotes the law of the solution X_t, while W is a Wiener process on an extended space, so that the natural filtration is extended with an initial independent sigma algebra $\mathcal{G}_0$, to make ξ an $\mathcal{F}_0$-measurable random variable with law μ_0. As shown by Gärtner, under appropriate conditions on the coefficients, there exists a unique strong solution of (1.2), X_t, and its law, μ_t satisfies (1.1).

As we mentioned before, sde (1.2) is sometimes called non-linear, since its coefficients involve at the same time X_s and its law. In [BT1], it is suggested that the numerical approximation of (1.2) must act on two levels. On one, the usual time discretization (see [KP]) is needed, based on simulations of the increments of the driving process W. On the other, it is necessary to use some empirical measure in order to approximate the measures μ_s that appear in the coefficients. To this purpose, the simulation scheme is expanded introducing n independent driving Wiener processes, each generating a particle through an equation that approximates (1.2) (for details see Section 3). These particles, denoted by X^i, $i = 1, ..., n$, will interact with each other through their empirical measure, viewed as an approximation of μ_s. By some kind of law of large numbers (or propagation of chaos as it is better known), this interaction tends to disappear as $n \to \infty$.

Bossy and Talay prove that the empirical distribution generated by the X^i converges to the law of X and therefore give a method to approximate the solution of the McKean -Vlasov equation (1.1). More exactly they prove the following result, which we report here for the reader's convenience, since we will comparatively refer to it.

Theorem 1.1: *Let $a(x, y) = b(x, y) = y$ and assume*

(H-1) there exists a strictly positive constant c such that $g(x, y) \geq c > 0$, $\forall (x, y) \in \mathbb{R}^2$;
(H-2) the functions f and g are uniformly bounded on $\mathbb{R}^2$; f is globally Lipschitz and g has uniformly bounded first partial derivatives;
(H-3) the initial law μ_0 satisfies one of the following
(i) μ_0 is a Dirac measure at x_0
(ii) μ_0 has a continuous density p_0 so that there exist constants $M, \alpha > 0$, $\eta \geq 0$ such that $p_0(x) \leq \eta \exp(-\alpha \frac{x^2}{2})$ for $|x| > M$ (if $\eta = 0$, μ_0 has compact support).

Furthermore, if $u(t, \cdot)$ is the distribution function of X_t and $\overline{u}(t, \cdot)$ the empirical distribution function of the sequence X_t^i for $i = 1, ..., n$, then for any fixed $t \in [0, T]$

$$E\|u(t, \cdot) - \overline{u}(t, \cdot)\|_{L^1(\mathbb{R})} \leq C(\frac{1}{\sqrt{n}} + \sqrt{h}). \tag{1.3}$$

If we substitute (H-2) and (H-3) with the stronger conditions

(H-2') $f \in C_b^2(\mathbb{R}^2)$ and $g \in C_b^3(\mathbb{R}^2)$.

(H-3') The initial law μ_0 has a strictly positive density $p_0 \in C^2(\mathbb{R})$ and there exist constants $M, \eta, \alpha > 0$ such that $p_0 + |p_0'(x)| + |p_0''(x)| \leq \eta \exp(-\alpha \frac{x^2}{2})$ for $|x| > M$,

then μ_t has a density, denoted by $p_t(\cdot)$, and

$$E\|p_t(\cdot) - \frac{1}{n} \sum_{j=1}^{n} \phi_\epsilon(X_t^j - x)\|_{L^1(\mathbb{R})} \leq C(\epsilon + \frac{1}{\sqrt{\epsilon}}(\frac{1}{\sqrt{n}} + \sqrt{h})), \tag{1.4}$$

where $\phi_\epsilon(z) = \frac{e^{-\frac{z^2}{2\epsilon}}}{\sqrt{2\pi\epsilon}}$.

The goal of our work is to prove that the rate in (1.3) is actually $\frac{1}{\sqrt{n}} + h$ under conditions comparable to (H-1), (H-2) and (H-3). We will first establish the result for the densities showing that the optimal rate in (1.4) is of the order $\frac{1}{\sqrt{nh}} + h$, when $\epsilon = h$, rather than $h + \frac{1}{\sqrt{nh}} + \sqrt{h} + 1$.

Our efforts clearly drew inspiration from the remarks made by Bossy and Talay (see [BT1] and [BT2]), who gave numerical evidence that suggested the rate of convergence was faster than what they proved.

Here we are able to achieve this better rate, by using completely different techniques from those in [BT1]. Indeed, we carefully employ Malliavin Calculus techniques together with some ideas brought to light in a recent work by Kohatsu and Ogawa ([KO]).

Malliavin Calculus allows to establish when the marginal densities of the solution of a sde exist and are regular, so it is indeed very apt to deal with equations, whose coefficients involve probability densities. The introduction of these techniques in this setting enabled us also to weaken slightly the hypotheses on the coefficients as well as those on the initial density function. We establish this result in Section 2 and it is adapted from the similar one obtained by Taniguchi (see [T]), in the study of the smoothness of densities for time dependent systems.

The key idea in the proof of the result is that the coefficients have to verify the so called "restricted" Hörmander condition. The main difference between Taniguchi's results and ours is that we do not require any boundedness for the coefficients, but bounded differentiability in all the required derivatives is indeed sufficient. This property is in fact satisfied by the coefficients of (1.2), so we can apply the same results of existence and smoothness of the densities to the process under study. Another difference is the introduction of an initial random variable. If we were to use a uniform restricted Hörmander type condition, like in [T], this difference would be minor. But applications force the study of the case when the initial random variable is supported on the whole real line. Therefore such a uniform restricted Hörmander condition would be very restrictive. Here we only require some tail conditions on the initial random variable. In order to carry out the proof in this case one needs to study carefully the behaviour of all the bounds with respect to the initial random variable.

In Section 3 we study the approximation errors of the particle method used to approximate the solution of (1.2), this analysis relies on a technique very different from the one used by Bossy and Talay. We try to avoid as much as possible any L^p estimates in order to obtain the rate h instead of $\sqrt{h}$ in (1.3). This is obtained via an approximation method which is briefly explained at the end of Theorem 3.7.

The basic idea is as follows: consider formally the quantity

$$
\begin{aligned}
&E\|E(\delta_x(X_t)) - \frac{1}{n}\sum_{j=1}^{n}\phi_\epsilon(X_t^j - x)\|_{L^1(\mathbb{R})} \\
&\qquad \leq \|E(\delta_x(X_t)) - E(\delta_x(X_t^1))\|_{L^1(\mathbb{R})} + E\|E(\delta_x(X_t^1)) - \frac{1}{n}\sum_{j=1}^{n}\phi_\epsilon(X_t^j - x)\|_{L^1(\mathbb{R})}
\end{aligned}
$$

The second term is about the order $\frac{1}{\sqrt{n}}$ (some correlation structure between the X^j has to be studied). The first is a term of the same kind that arises in classical weak approximation procedures, except that in our case discretization both in time and in space (measure discretization) is used. By analyzing separately the two discretizations one gets a rate of convergence of the order of $h + \frac{1}{\sqrt{n}}$.

To carry out this idea is not as easy as explained above. It presents some extra complications with respect to the classical case of diffusions and it is essential for our method to work, that we run a separate study of the time and space discretizations.

The results for approximations of the distribution function of X_t are obtained with similar techniques as those used for the density functions. For this reason we decided to explain in detail this second case, technically more demanding, and to sketch the proofs for the first.

We hope the methods exposed here will help develop similar results also for the Burgers equation and in general, for non-linear equations.

We should note, as remarked by Bossy and Talay, that this is not the only method available to simulate the solution to the McKean-Vlasov equation. There is also the random vortex methods developed by Chorin and Hald (see [C], [H1], [H2]) which are based on a splitting of the nonlinear differential operator.

The paper is subdivided as follows. In Section 2, we give the preliminary results that enable to conclude the existence and smoothness of the densities of the solution of (1.2) under restricted Hörmander conditions. Given that the problem in one dimension simplifies to elliptic conditions we have decided to consider the multidimensional case. This also allows us to present a counterexample to the general unrestricted Hörmander condition.

In the rest of the paper we concentrate on the one dimensional case, the general case being a straightforward generalization. In Section 3 we establish our approximation results for densities, while in section 4 we summarize those and we derive the distribution function case.

As usual we adopt the convention of writing the same letter (usually C) for a constant even if it changes from line to line. This constant is always independent of h, n and the partition of the time interval. Unless otherwise stated we will also assume without loss of generality that all constants are bigger than 1.

2. Preliminary results

Let $[0,T]$ be a finite time interval and $(\Omega, \mathcal{F}, P)$ a complete probability space, where a standard $d-$dimensional Brownian motion, W, is defined. We consider the equation in $\mathbb{R}^n$

$$X_t = \xi + \int_0^t a(X_s, F(X_s;\mu_s))ds + \int_0^t b(X_s, G(X_s;\mu_s)) \cdot dW_s, \tag{2.1}$$

where ξ is an $\mathcal{F}_0$-measurable random variable in $\mathbb{R}^n$, such that $\xi \in \bigcap_{p\geq 1} L^p$. By $F(x;\mu_s)$ or $G(x;\mu_s)$ we denote the functions given by $\int_{\mathbb{R}^n} \zeta(x,y)d\mu_s(y)$ ($\zeta = f$, g, respectively), where μ_s indicates the distribution of X_s. Lastly, the functions

$$a\,:\mathbb{R}^n\times\mathbb{R}\longrightarrow\mathbb{R}^n, \qquad b\,:\mathbb{R}^n\times\mathbb{R}\longrightarrow\mathbb{R}^n\times\mathbb{R}^d, \qquad f,\,g\,:\mathbb{R}^n\times\mathbb{R}^n\longrightarrow\mathbb{R}$$

are all smooth with bounded derivatives, let us call M the common constant dominating these all. This set of hypotheses will hold throughout the paper and we refer to it as (H0).

We are going to study the existence and smoothness of the density of the solution of (2.1). From now on, for ease of writing, we will call $\bar{a}(t,x) = a(x, F(x;\mu_t))$ and $\bar{b}(t,x) = b(x, G(x;\mu_t))$. Next, we introduce a series of hypotheses that we need for our goal.

First of all, let us remark that equation (2.1) maybe rewritten as

$$X_t = \xi + \int_0^t \bar{a}(s, X_s)ds + \sum_{k=1}^{d} \int_0^t \bar{b}_k(s, X_s)dW_s^k, \tag{2.2}$$

where we explicitly wrote the $d-$dimensional Brownian motion.

Assumptions:

(H1) There exists an integer m_0 and a positive constant c, such that

$$\sum_{i=0}^{m_0} \sum_{v \in I_i} < v(0, \xi), \eta >^2 \geq c > 0 \quad \text{a.e.} \quad \text{for all } \eta \in S^{n-1}$$

where the sets I_i are given by

$$I_0 = \{\bar{b}_1, \dots, \bar{b}_d\}, \ \dots \ , \ I_n = \{ \ [\bar{b}_k, v_h], \ v_h \in I_{n-1}, \ 1 \leq h, k \leq d\}$$

and ($[\cdot, \cdot]$ denotes the Lie bracket. In this context, the coefficients are to be understood as vector fields, that is $\bar{b}_k(t,x) = \sum_{i=1}^{n} \bar{b}_k^i(t,x) \frac{\partial}{\partial x_i}$).

(H2) The functions b_k are bounded, let us say by the same constant M as in (H0).

(H3) ξ has a density u_0 for which there exist positive constants η, α, β and ρ such that

$$u_0(x) \leq \eta \exp(-\alpha |x|^\beta) \quad \text{for } \ |x| \geq \rho.$$

Hypothesis (H1) is the so called "restricted" Hörmander condition, as it involves only the diffusion coefficients. In the one dimensional case this reduces to saying that almost surely $|\bar{b}(0, \xi)| \geq c > 0$, which is very similar to the corresponding (H-1) in Theorem 1.1, requiring $\bar{b}(0, x) \geq c > 0$, for all $x \in \mathbb{R}$. In the multidimensional case, instead (H1) becomes actually much weaker than (H-1), as it may involve the brackets of order higher than one, while (H-1) does not.

Hypothesis (H2) is similar to (H-2') in Theorem 1.1, note that the smoothness in the coefficients is needed here to be able to study the smoothness of the density. Finally, (H3) is slightly weaker than the corresponding (H-3').

Another difference is given by the fact that in Theorem 1.1 all three conditions are assumed, while we are going to show, by means of Malliavin Calculus techniques, that is necessary to assume only (H1) and either (H2) or (H3). Therefore the combination of (H1) and (H2) give that the restricted Hörmander condition does not have to be necessarily uniform as is required in (H-1).

Since all the results in the paper rely heavily on Malliavin Calculus, we want to introduce here some of its terminology very briefly.

For $d \in \mathbb{N}$, we denote by $C_b^\infty(\mathbb{R}^d)$ the set of C^∞ bounded functions $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, with bounded derivatives of all orders and we assume that a $d-$dimensional Wiener process is defined on a probability space, (actually we will use $d = 1,\ d = 2$ or $d = 3$).

If we denote by S the class of real random variables F that can be represented as $f(W_{t_1}, \dots, W_{t_n})$ for some $n \in \mathbb{N}$, $t_1, ..., t_n \in [0, T]$ and $f \in C_b^\infty(\mathbb{R}^{nd})$, we can complete this space under the norm $\| \cdot \|_{1,p}$ given by

$$\| F \|_{1,p}^p = E(|F|^p) + \left(\sum_{j=1}^d E(\int_0^T |D_s^j F|^2 ds)^{\frac{p}{2}} \right),$$

where D^j is defined as $D_s^j F = \displaystyle\sum_{i=1}^n \frac{\partial f}{\partial x_{ij}}(W_{t_1}, \dots, W_{t_n}) 1_{[0,t_i]}(s)$, for $j = 1, \dots, d$, obtaining a Banach space, usually indicated with $\mathbb{D}^{1,p}$. Analogously, we can construct the space $\mathbb{D}^{k,p}$, by completing S under the norm

$$\| F \|_{k,p}^p = E(|F|^p) + \sum_{j=1}^k \sum_{k_1+\dots+k_d=j} E\left((\int_0^T \dots \int_0^T |D^{d,k_d}_{s_j \dots s_{j-k_d}} \dots D^{1,k_1}_{s_{k_1} \dots s_1} F|^2 ds_1 \dots ds_j)^{\frac{p}{2}} \right),$$

where $D^{i,l}_{s_1 \dots s_l} F = D^i_{s_1} \dots D^i_{s_l} F$. Finally, we denote $\mathbb{D}^\infty = \displaystyle\bigcap_{p\geq 1} \bigcap_{k \geq 1} \mathbb{D}^{k,p}$.

The adjoint of the closable unbounded operator

$$D^j : \mathbb{D}^{1,2} \subseteq L^2(\Omega) \longrightarrow L^2([0, T] \times \Omega)$$

is usually denoted by δ^j and it is called the Skorohod integral. The domain of δ^j is the set of all processes $u \in L^2([0, T] \times \Omega)$ such that

$$\left| E \left(\int_0^T D_t^j F u_t dt \right) \right| \leq C \| F \|_2 \qquad \forall\, F \in S,$$

for some constant C depending possibly on u.

If $u \in Dom(\delta^j)$, then $\delta^j(u)$ is the square integrable random variable determined by the duality relation

$$E(\delta^j(u) F) = E(\int_0^T D_t^j F u_t dt) \qquad \forall\, F \in \mathbb{D}^{1,2}.$$

In the multidimensional case we consider $\delta = \sum_j \delta^j$.

Finally, for a, possibly m-dimensional, random variable F we denote its Malliavin covariance matrix by γ_F and it is defined as

$$\gamma_F^{hk} = \sum_{j=1}^d \int_0^T D_s^j F^h\, D_s^j F^k ds \qquad h, k = 1, ..., m.$$

The Malliavin covariance matrix plays a key role when one wants to determine the existence and the smoothness of the densities of the solutions of stochastic differential equations. Namely, following [N1] (Proposition 2.1.1, page 78), we have that for any random variable $F \in \mathbb{D}^{1,p}_{loc}$ for some $p > 1$, if γ_F is almost surely invertible, then the law of F is absolutely continuous with respect to Lebesgue measure, moreover if $F \in \mathbb{D}^{1,2}$ (and it is one dimensional) and $\gamma_F^{-1} DF$ is in $Dom(\delta)$ then F has a continuous and bounded density given by

$$f(x) = E\left(1_{\{F>x\}}\delta(\gamma_F^{-1}DF)\right).$$

In particular we will use the fact that if $F \in \mathbb{D}^{\infty}$ and $|\gamma_F^{-1}| \in \bigcap_{p>1} L^p$ then F has an infinitely differentiable density (see [N1], Corollary 2.1.2).

Remark : We would like to point out that the restricted Hörmander condition is indeed needed for the densities to be absolutely continuous with respect to Lebesgue measure, when considering time dependent coefficients. In fact it is possible to construct easy examples where the coefficients of a SDE verify an unrestricted Hörmander condition at all times and points, but for which the Malliavin covariance matrix associated to the solution is not invertible (see [T]).

Having introduced all the necessary terminology we first quote a result from [KO] about existence and integrability of the solution of (2.1).

Theorem 2.1: *Let us assume that (H0) is satisfied, then there is a unique strong solution of (2.1) such that, for all* $p > 1$

$$E(\sup_{s\le T} |X_s|^p) \le \infty.$$

Furthermore $X_s \in \mathbb{D}^{\infty}$ *for all* $s \in [0,T]$.

We are now able to state and prove the main result of this section about the marginal densities of $X_.$.

Theorem 2.2: *Assume that (H0) and (H1) are satisfied together with either (H2) or (H3). Then* $\gamma_{X_t}^{-1} \in \bigcap_{p\ge1} L^p$ *and* X_t *has a smooth density.*

PROOF: Our proof is an adaptation of the method used by [T], extended to the case when the initial point is random and the uniformity on the Hörmander condition is relaxed. Since the argument is basically the same under either (H2) or (H3), we will prove the result assuming the latter, being the more difficult one, and we will point out the differences with the other case step by step.

As a starting point, let us remark that in equation (2.2) the coefficients are time dependent. Moreover, because of (H0), $\bar{a}$ and $\bar{b}_k$ are smooth in space with bounded derivatives (hence they are also globally Lipschitz) and they are globally differentiable in time as many times as needed. Indeed, for $,t \in [0,T]$, we have for any k,

$$\frac{\partial \bar{b}_k}{\partial t}(t,x) = \frac{\partial b_k}{\partial t}(x, E(g(x,X_t))) = \frac{\partial b_k}{\partial y}(t, E(g(x,X_t)))\frac{\partial E(g(x,X_t))}{\partial t}.$$

On the other hand, applying Itô's Lemma to $g(x,\cdot)$, we get

$$\begin{aligned} E(g(x,X_t)) = E(g(x,\xi)) + \int_0^t E[\sum_{i=1}^n g_{x_i}(x,X_r)\bar{a}^i(r,X_r)]dr \\ + \int_0^t E[\sum_{i,j=1}^n g_{x_ix_j}(x,X_r)\sum_{k=1}^d \bar{b}_k^i(r,X_r)\bar{b}_k^j(r,X_r)]dr \end{aligned}$$

and hence we obtain that

$$\frac{\partial E(g(x,X_t))}{\partial t} = E[\sum_{i=1}^n g_{x_i}(x,X_t)\bar{a}^i(t,X_t) + \sum_{i,j=1}^n g_{x_ix_j}(x,X_t)\sum_{k=1}^d \bar{b}_k^i(t,X_t)\bar{b}_k^j(t,X_t)]$$

which is bounded. In fact all the derivatives of the coefficient g are bounded by M by hypothesis, besides we can bound $E(|\bar{a}^i(r,X_r)|)$ and $E(|\bar{b}_k^i(t,X_t)\bar{b}_k^j(t,X_t)|)$, by exploiting the mean value theorem and hypothesis (H0). For instance, for $E(|\bar{a}^i(r,X_r)|)$ we have

$$\begin{aligned} |\bar{a}^i(r,X_r)| &= |a^i(X_r, \int f(X_r,y)\mu_r(dy))| \\ &= |a^i(X_r, \int_{\mathbb{R}^n} f(X_r,y)\mu_r(dy)) - a^i(\xi, \int_{\mathbb{R}^n} f(\xi,y)\mu_0(dy))| + |a^i(\xi, \int f(\xi,y)\mu_0(dy))| \\ &\leq M(|X_r-\xi| + |\int_{\mathbb{R}^n} f(X_r,y)\mu_r(dy)) - \int_{\mathbb{R}^n} f(\xi,y)\mu_0(dy))|) \\ &+ |a^i(\xi, \int f(\xi,y)\mu_0(dy)) - a^i(0, \int f(0,y)\mu_0(dy))| + |a^i(0, \int f(0,y)\mu_0(dy))| \\ &\leq M(|X_r-\xi| + M|X_r-\xi| + |\int_{\mathbb{R}^n} f(\xi,y)\mu_r(dy) - \int_{\mathbb{R}^n} f(\xi,y)\mu_0(dy)|) \\ &+ M(|\xi| + \int_{\mathbb{R}^n} |f(\xi,y)-f(0,y)|\mu_0(dy)) + |a^i(0, \int f(0,y)\mu_0(dy))|) \\ &\leq M(|X_r-\xi| + 2M|X_r-\xi|) + M(|\xi| + M|\xi| + |a^i(0, \int f(0,y)\mu_0(dy))|). \end{aligned}$$

Taking expectations, the last side is finite. The same argument applies to the other term.

Of course, the same procedure may be repeated for $\bar{a}$ and the spatial and time derivatives of any order, proving the smoothness in time of the coefficients of the equation. It is exactly this smoothness that allows us, differently from Taniguchi' paper, to include in our discussion coefficients that might be not bounded.

Let ζ_t denote the derivative of the stochastic flow associated with equation (2.2) and ζ^{-1} the inverse flow, then both $\sup_{s\le t}|\zeta_t(\xi)|$ and $\sup_{s\le t}|\zeta_t^{-1}(\xi)| \in \bigcap_{p\ge 1} L^p$ and we can write the Malliavin derivative as $D_s^k X_t = \zeta_t(\xi)\zeta_s^{-1}(\xi)\bar{b}_k(s,X_s)$ (see [N1], page 109), where by D^k we are denoting the Malliavin derivative with respect to the $k-$th component of the Brownian motion.

Since we already noticed that $X_t \in \mathbb{D}^\infty$, to conclude the existence and regularity of the density, it is enough to check that

$$\gamma_{X_t}^{-1} = \sum_{k=1}^d [\zeta_t(\xi)\int_0^t \zeta_s^{-1}(\xi)\bar{b}_k(s,X_s)\bar{b}_k(s,X_s)^T\zeta_s^{-1}(\xi)^T ds\, \zeta_t(\xi)^T]^{-1} \in \bigcap_{p\ge 1} L^p,$$

where by T we denote the transpose. Using the L^p boundedness of $\zeta_t(\xi)$ and Lemma 2.3.1 in [N1], this amounts to show that for all $p\ge 2$ there exists $\epsilon_0(p)$ such that for all $\epsilon \le \epsilon_0(p)$

$$P\left(\sum_{k=1}^d \int_0^t \zeta_s^{-1}(\xi)\bar{b}_k(s,X_s)\bar{b}_k(s,X_s)^T\zeta_s^{-1}(\xi)^T ds < \epsilon\right) \le \epsilon^p. \tag{2.3}$$

Following [T] or [N1], to prove (2.3), it suffices to show that for all η in the unit ball S^{n-1}, one has

$$P\left(\sum_{k=1}^d \int_0^t <\zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta>^2 ds < \epsilon\right) \le \epsilon^p. \tag{2.4}$$

In particular we will prove that for the same m_0 of hypothesis (H1), some function ν and any $N\in\mathbb{N}$, we have

$$P\left(\sum_{k=1}^d \int_0^t <\zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta>^2 ds < \frac{1}{N^{\nu(m_0)}}\right) \le \frac{1}{N^{p\nu(m_0)}}. \tag{2.5}$$

In order to do so, we divide this probability into two parts, by fixing $\tau\in\mathbb{R}^+$

$$\begin{aligned}
&P\left(\sum_{k=1}^d \int_0^t <\zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta>^2 ds < \epsilon\right)\\
\le &P\left(\sum_{k=1}^d \int_0^t <\zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta>^2 ds < \epsilon,\ |\xi|\le\tau\right) + P(|\xi|\ge\tau)\\
= &p_1 + p_2.
\end{aligned}$$

Clearly for p_2 we will use hypothesis (H3), which gives $P(|\xi|>\tau)\le C\exp(-\alpha\tau^\beta)$ for $\tau\ge\rho$. At the end of the proof, we will specify how to choose τ to have (2.3) satisfied.

As for p_1, we first note that, given (H1) and $|\xi| \le \tau$, one also has that for all $\eta \in S^{n-1}$,

$$\sum_{i=0}^{m_0} \sum_{v \in I_i} < v(s,y), \eta >^2 \ge \frac{3c}{4} > 0, \tag{2.6}$$

for $|y - \xi| \le R$ and $s \le R$, where we can chose

$$R = \frac{\sqrt{c}}{32(2M)^{m_0} n^{m_0 - 1}(4+\tau)}.$$

where M dominates all derivatives and $\max\limits_{\substack{1 \le k \le d \\ 1 \le i \le n}} |b_k^i(0,0)|$. Moreover

$$R \in [\frac{\sqrt{c}}{L(4+\tau+\sqrt{c})}, \frac{\sqrt{c}}{4L}], \tag{2.7}$$

where $L = 32(2M)^{m_0} n^{m_0-1}$.

We remark that if we are instead assuming (H2), it is not necessary to split the probability in (2.4) into two parts and besides (2.6) holds automatically with

$$R \le \frac{\sqrt{c}}{64(2M)^{m_0} n^{m_0-1}}. \tag{2.7'}$$

Let us define the stopping time

$$\sigma_N = \inf\{s : \ |X_s - \xi| \ge \frac{R}{2} \text{ or } |\zeta_s^{-1}(\xi) - 1| \ge \frac{1}{2}\} \wedge \frac{R}{2} \wedge \frac{1}{N^3}.$$

Given t, for N sufficently large we may assume that $\frac{1}{N^3} \le \min(\frac{R}{2}, t)$ hence $\sigma_N \le \frac{1}{N^3}$ and

$$p_1 \le P(\sum_{k=1}^{d} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi)\bar{b}_k(s, X_s), \eta >^2 \, ds < \epsilon, |\xi| \le \tau). \tag{2.8}$$

Let us remark that by definition of σ_N, for any $q > 1$, we have

$$\begin{aligned}
&P(\sigma_N < \frac{1}{N^3}) \le P(\sup_{0 \le s \le \frac{1}{N^3}} |X_s - \xi| \ge \frac{R}{2}) + P(\sup_{0 \le s \le \frac{1}{N^3}} |\zeta_s^{-1}(\xi) - 1| \ge \frac{1}{2}) \\
&\le (\frac{2}{R})^q E(\sup_{0 \le s \le \frac{1}{N^3}} |X_s - \xi|^q) + 2^q E(\sup_{0 \le s \le \frac{1}{N^3}} |\zeta_s^{-1}(\xi) - 1|^q) \\
&\le \frac{C_1 2^q}{R^q} \left(\frac{1}{N^3}\right)^{q/2} + C_2 2^q \left(\frac{1}{N^3}\right)^{q/2},
\end{aligned}$$

for some constants C_1 and C_2.

At this point, let us remark that

$$\sum_{k=1}^{d} < \zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta >^2 = \sum_{v\in I_0} < \zeta_s^{-1}(\xi)v(s,X_s),\eta >^2$$

and let us introduce, as it is done in [T], the following sets

$$E_j(N) = \{\sum_{i=0}^{j}\sum_{v\in I_i}\int_0^{\sigma_N} < \zeta_s^{-1}(\xi)v(s,X_s),\eta >^2 \, ds < \frac{d^j}{N^{\gamma_j}}\} \cap \{|\xi| \le \tau\}$$

where $\gamma_j = 5\cdot 4^{m_0-j}$, for $j = 1,\dots,m_0$ (note that $\gamma_j = 4\gamma_{j+1}$).

We now chose in (2.4) $\epsilon = \frac{1}{N^{\gamma_0}}$, that is to say that $\nu(m_0) = 5\cdot 4^{m_0}$, therefore from (2.8) we get

$$\begin{aligned}
p_1 \le & P(\sum_{v\in I_0}\int_0^{\sigma_N} < \zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta >^2 \, ds < \frac{1}{N^{\gamma_0}}, |\xi| \le \tau) \\
= & P(E_0(N)) \le P(E_0(N) \cap \{\sigma_N < \frac{1}{N^3}\}) + P(E_0(N) \cap \{\sigma_N = \frac{1}{N^3}\}) \\
\le & P(\sigma_N < \frac{1}{N^3}) + \sum_{j=0}^{m_0-1} P(E_j(N) \cap E_{j+1}(N)^c \cap \{\sigma_N = \frac{1}{N^3}\}) + P(E_{m_0}(N) \cap \{\sigma_N = \frac{1}{N^3}\})
\end{aligned}$$

The last probability is actually zero, because of the Hörmander condition (2.6), indeed on the set $\{\sigma_N = \frac{1}{N^3}\}$, we necessarily have that, on $\{s \le \sigma_N\}$, $\dfrac{1}{2} \le |\zeta_s^{-1}(\xi)| \le \dfrac{3}{2}$ and $|X_s - \xi| \le \frac{R}{2}$, hence (2.6) holds a.s. with X_s in place of y and consequently

$$\sum_{i=0}^{m_0}\sum_{v\in I_i}\int_0^{\sigma_N} < \zeta_s^{-1}(\xi)v(s,X_s),\eta >^2 \, ds \ge \frac{3}{8}c\sigma_N,$$

for an appropriate positive constant c. This implies that on the set

$$E_{m_0}(N) \cap \{\sigma_N = \frac{1}{N^3}\} = \{\sum_{i=0}^{m_0}\sum_{v\in I_i}\int_0^{\sigma_N} < \zeta_s^{-1}(\xi)v(s,X_s),\eta >^2 \, ds \le \frac{d^{m_0}}{N^5}, \sigma_N = \frac{1}{N^3}, |\xi| < \tau\}$$

we have

$$\frac{3}{8}\frac{c}{N^3} \le \sum_{i=0}^{m_0}\sum_{v\in I_i}\int_0^{\sigma_N} < \zeta_s^{-1}(\xi)v(s,X_s),\eta >^2 \, ds \le \frac{d^{m_0}}{N^5}$$

which is clearly not verified, as soon as N is large enough.

It remains to analyze the probabilities $P(E_j(N) \cap E_{j+1}(N)^c \cap \{\sigma_N = \frac{1}{N^3}\})$, $j = 0, \ldots, m_0$. On this set, we have

$$\sum_{i=0}^{j} \sum_{v \in I_i} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds \leq \frac{d^j}{N^{\gamma_j}} \quad \text{and}$$
$$\sum_{i=0}^{j+1} \sum_{v \in I_i} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds > \frac{d^{j+1}}{N^{\gamma_{j+1}}}.$$

Under these conditions we obtain

$$\begin{aligned}
&\sum_{i=0}^{j+1} \sum_{v \in I_i} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds \geq \sum_{v \in I_{i+1}} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds \\
=&\sum_{i=0}^{j+1} \sum_{v \in I_i} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds - \sum_{i=0}^{j} \sum_{v \in I_i} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds \\
\geq&\frac{d^{j+1}}{N^{\gamma_{j+1}}} - \frac{d^j}{N^{\gamma_j}} \geq \frac{d^{j+1}}{2N^{\gamma_{j+1}}}
\end{aligned}$$

On the other hand, the cardinality of the set $\{v \in I_i,\ i = 1, \ldots, j\}$ is less than $2d^j$, for any $j = 0, \ldots, m_0$, thus at least one of the terms in the above sum must be greater than or equal to $\frac{d}{2m_0 N^{\gamma_{j+1}}}$, moreover

$$\frac{d^j}{N^{\gamma_j}} = \frac{d^j}{N^{4\gamma_{j+1} - 9 + 9}} \leq \frac{1}{N^{4\gamma_{j+1} - 9}},$$

if $N^9 > d^{m_0} > d^j$, for all j. Taking these remarks into account, we may conclude that

$$\begin{aligned}
E_j(N) \cap E_{j+1}(N)^c \subseteq \bigcup_{i=0}^{j} \bigcup_{v \in I_i} \{ &\int_0^{\sigma_N} < \zeta_s^{-1}(\xi) v(s, X_s), \eta >^2 \, ds \leq \frac{1}{N^{4\gamma_{j+1} - 9}} \}, \\
&\cap \{ \sum_{k=1}^{d} \int_0^{\sigma_N} < \zeta_s^{-1}(\xi) [\bar{b}_k, v](s, X_s), \eta >^2 \, ds > \frac{d}{2m_0 N^{\gamma_{j+1}}} \} \\
&\cap \{ |\xi| \leq \tau \}.
\end{aligned}$$

As we remarked before, the coefficients of equation (2.2) are differentiable in time and space as many times as needed and consequently so are the vector fields $v \in I_i$, $i = 1, \ldots m_0$. thus, applying Itô's lemma and the integration by parts to $\zeta_s^{-1}(\xi) v(s, X_s)$ we obtain

$$\begin{aligned}
\zeta_s^{-1}(\xi) v(s, X_s) =& v(0, \xi) + \int_0^s \zeta_r^{-1}(\xi) \{ \frac{\partial v}{\partial s} + [\bar{a}, v] + \frac{1}{2} \sum_{h=1}^{d} [\bar{b}_h, [\bar{b}_h, v]] \}(r, X_r) dr \\
&+ \sum_{h=1}^{d} \int_0^s \zeta_r^{-1}(\xi) [\bar{b}_h, v](r, X_r) dW_r^h
\end{aligned}$$

that implies that $\sum_{h=1}^{d}\zeta_r^{-1}(\xi)[\bar{b}_h,v](r,X_r)$ is the diffusion coefficient of $\zeta_s^{-1}(\xi)v(s,X_s)$.

With considerations similar to those employed before, it is possible to show that when $s\le\sigma_N$ and $|\xi|\le\tau$, we have

$$\begin{aligned}&|\zeta_s^{-1}(\xi)|\le\frac{3}{2}\\ &|\bar{b}_k(s,X_s)|\le M(s+|X_s-\xi|+|\xi|+|b(0,0)|)\le M(2R+\tau+M)\le M(\frac{\sqrt{c}}{2L}+\tau+M)\end{aligned}\tag{2.9}$$

by the properties of R. Similarly we may conclude that there exists a constant λ, depending on n,M,m_0,c,d, so that for $s\le\sigma_N$ and $|\xi|\le\tau$

$$|\zeta_s^{-1}(\xi)v(s,X_s)|,\qquad \sum_{h=1}^{d}\zeta_s^{-1}(\xi)[\bar{b}_h,v](s,X_s)\le\lambda(1+\tau)\tag{2.10}$$

for all $v\in I_i$, with $i=1,\dots,m_0$.

We would like to remark that if we assume (H2) instead of (H3), this whole argument goes through, with the only difference that estimate (2.10) is valid independently of τ.

As it is shown in [T], if we apply Theorem 8.26 of [S] , we obtain for $m=\gamma_{j+1}\ge 5$,

$$\begin{aligned}&P(E_j(N)\cap E_{j+1}(N)^c,\sigma_N=\frac{1}{N^3})\\ \le&\sum_{i=0}^{j}\sum_{v\in I_i}P\left(|\xi|\le\tau,\sigma_N=\frac{1}{N^3},\int_0^{\sigma_N}<\zeta_s^{-1}(\xi)v(s,X_s),\eta>^2\,ds\le\frac{1}{N^{4\gamma_{j+1}-9}},\right.\\ &\left.\qquad\sum_{k=1}^{d}\int_0^{\sigma_N}<\zeta_s^{-1}(\xi)[\bar{b}_k,v](s,X_s),\eta>^2\,ds>\frac{d}{2m_0N^{\gamma_{j+1}}}\right)\\ \le&\sum_{i=0}^{j}\sum_{v\in I_i}\sqrt{2}N^{\gamma_{j+1}-5}\exp\{-\frac{dN}{2^9m_0^2\lambda^2(1+\tau)^2}\}.\end{aligned}$$

At last, summarizing, we can conclude

$$\begin{aligned}&P\left(\sum_{k=1}^{d}\int_0^t<\zeta_s^{-1}(\xi)\bar{b}_k(s,X_s),\eta>^2\,ds<\frac{1}{N^{5\cdot4^{m_0}}}\right)\\ \le&P(|\xi|\ge\tau)+P(\sigma_N<\frac{1}{N^3})+\sum_{j=0}^{m_0-1}P(E_j(N)\cap E_{j+1}(N)^c\cap\{\sigma_N=\frac{1}{N^3}\})\\ \le&Ce^{-\alpha\tau^\beta}+(\frac{C_12^q}{R^q}+C_22^q)\left(\frac{1}{N^3}\right)^{q/2}+\sum_{j=0}^{m_0-1}\sum_{i=0}^{j}\sum_{v\in I_i}\sqrt{2}N^{\gamma_{j+1}-5}\exp\{-\frac{dN}{2^9m_0^2\lambda^2(1+\tau)^2}\}\\ \le&Ce^{-\alpha\tau^\beta}+(\frac{C_12^q}{R^q}+C_22^q)\left(\frac{1}{N^3}\right)^{q/2}+\sum_{j=0}^{m_0-1}m_0d^j\sqrt{2}N^{\gamma_{j+1}-5}\exp\{-\frac{dN}{2^9m_0^2\lambda^2(1+\tau)^2}\}\end{aligned}$$

and the result follows by taking $\tau=O(|log((\frac{1}{N^3})^{q/\alpha})|^{1/\beta})$. $\square$

We would like to remark that it is exactly the use of the restricted Hörmander condition that makes the above theorem true for time dependent coefficients, since it enables us to use Theorem 8.26 of [S], which involves only the diffusion coefficients of the process and therefore does not call for an unrestricted condition that would require also the drift coefficient. In fact, one can find a counterexample to the general statement of existence of densities under the general Hörmander condition (see page 310 in [T]).

From the previous theorem we know that there exists a unique solution to (2.1) with smooth density, which we denote by $p_t(x)$, that we are eventually interested in approximating.

In order to relate the unique solution of (2.1) to the McKean-Vlasov equation we recall that, under appropriate conditions (see [G]), the distribution function of X_t, denoted by $u(t,x)$, satisfies the equation

$$(2.10)\qquad \begin{cases} \dfrac{\partial u}{\partial t}(t,x) = \dfrac{1}{2}\dfrac{\partial}{\partial x}[b^2(x, G(x; u(t,x)))\dfrac{\partial u}{\partial x}(t,x)] - a(x, F(x; u(t,x)))\dfrac{\partial u}{\partial x}(t,x) \\ u(0,x) = P(\xi \le x)\,. \end{cases}$$

Assuming enough regularity of the solution, by differentiating the above equation, one obtains that the density of X_t, denoted by $p_t(x) \equiv p(t,x)$, satisfies the following non-linear equation

$$(2.11)\qquad \begin{cases} \dfrac{\partial p}{\partial t}(t,x) = \dfrac{1}{2}\dfrac{\partial^2}{\partial x^2}[b^2(x, G(x; \int_{-\infty}^{x} p(t,y)dy))p(t,x)] - \dfrac{\partial}{\partial x}[p(t,x)a(x, F(x; \int_{-\infty}^{x} p(t,y)dy))] \\ u(0,x) = p_0(x)\,. \end{cases}$$

Therefore, it becomes of interest to approximate both the distribution and density function of X_t for fixed $t > 0$.

In order to do this, in the next section we introduce a particle method described in Bossy and Talay [BT1] and [BT2] and we evaluate the rate of convergence of this method to the solution.

3. Particle method

In this section we describe the actual particle method that we use to approximate $p_t(x)$, the density of the solution of (2.1).

In order to do so, we proceed by steps.

(1) Approximate the density $p_t(x)$ by Gaussian densities, i.e.

$$p_t(x) = \int_{\mathbb{R}} \delta_x(y) p_t(y) dy \sim \int_{\mathbb{R}} \phi_h(y-x) p_t(y) dy = E(\phi_h(X_t - x)),$$

where $\phi_h(z) = \frac{e^{-\frac{z^2}{2h}}}{\sqrt{2\pi h}}$.

(2) Consider the difference

$$p_t(x) - E(\phi_h(X_t - x)).$$

(3) Given a partition $\pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, which without loss of generality we assume to be uniform with mesh h, i.e. $h = \Delta t = t_{i+1} - t_i$ for any i, we define the Euler scheme for equation (2.1) as

$$Y_t = Y_{\eta(t)} + a(Y_{\eta(t)}, F(Y_{\eta(t)}; v_{\eta(t)}))(t - \eta(t)) + b(Y_{\eta(t)}, G(Y_{\eta(t)}; v_{\eta(t)}))(W_t - W_{\eta(t)}), \tag{3.1}$$

where $\eta(t) = \sup\{t_i \leq t \ : \ t_i \in \pi\}$ and $F(x; v_{\eta(t)}) = \int_{\mathbb{R}} f(x,y) dv_{\eta(t)}(y)$, with v_s denoting the distribution of Y_s.

(4) Consider the difference

$$E(\phi_h(X_t - x)) - E(\phi_h(Y_t - x))$$

(5) Generate n independent copies of the Euler scheme, that we denote by Y^i and consider the difference

$$E(\phi_h(Y_t - x)) - \frac{1}{n}\sum_{j=1}^{n} \phi_h(Y_t^j - x).$$

(6) Consider the Euler/weakly interacting particle system given by

$$X_t^i = X_{\eta(t)}^i + a(X_{\eta(t)}^i, F(X_{\eta(t)}^i; \bar{u}_{\eta(t)}))(t - \eta(t)) + b(X_{\eta(t)}^i, G(X_{\eta(t)}^i; \bar{u}_{\eta(t)}))(W_t^i - W_{\eta(t)}^i) \tag{3.2}$$

where $\bar{u}_{\eta(t)}(dx) = \frac{1}{n}\sum_{j=1}^{n} \delta_{X_{\eta(t)}^j}(dx)$.

(7) Consider the difference

$$\frac{1}{n}\sum_{j=1}^{n} \phi_h(Y_t^j - x) - \frac{1}{n}\sum_{j=1}^{n} \phi_h(X_t^j - x).$$

A similar procedure is followed to analyze the approximations for distributions functions, where the role of ϕ_h is played by its distribution function $\Phi_h(x) = \int_{-\infty}^{x} \phi_h(y) dy$. Our aim is to show the following result

Theorem 3.1: *Assume (H0), (H1) and either (H2) or (H3). Then for any fixed* $t \in (0,T]$,

$$\int_{\mathbb{R}} E\left(|u(t,x) - \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_t^j \leq x\}}|\right) dx \leq C(h + \frac{1}{\sqrt{n}}), \tag{3.3}$$

$$\int_{\mathbb{R}} E\left(|p_t(x) - \frac{1}{n}\sum_{j=1}^{n} \phi_h(X_t^j - x)|\right) dx \leq C(h + \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{nh}})\,. \tag{3.4}$$

Furthermore, if we choose $n = O(\frac{1}{h})^k$ *for some* $k > 0$*, then for each* $p > 1$*, there exists a positive constant* C_p *independent of* h *(and* n*) such that*

$$\sup_{x\in\mathbb{R}} E\left(|u(t,x) - \frac{1}{n}\sum_{j=1}^{n} \Phi_h(X_t^j - x)|\right) \leq C_p(h + \frac{1}{\sqrt{n}h^{\frac{p-1}{2}}}), \tag{3.5}$$

$$\sup_{x\in\mathbb{R}} E\left(|p_t(x) - \frac{1}{n}\sum_{j=1}^{n} \phi_h(X_t^j - x)|\right) \leq C_p(h + \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n}h^{1-\frac{1}{2p}}}). \tag{3.6}$$

Before moving toward this goal, we want to mention a result from [KO], that will provide an important requirement for the subsequent proofs.

Lemma 3.2 ([KO]): *Let* X_t *and* Y_t *be defined respectively by (2.1) and by (3.1) and let condition (H0) be fulfilled.*

Then $X_t, Y_t \in \mathbb{D}^\infty$ *for any* $t \in [0,T]$ *and for any* $n = 0, 1, \ldots$ *and any fixed* $q \geq 1$ *we have*

$$\sup_{s_1,\ldots,s_n \leq T} \| \sup_{t\leq T} |D_{s_1} \ldots D_{s_n} X_t \|_{2q} + \sup_{s_1,\ldots,s_n \leq T} \| \sup_{t\leq T} |D_{s_1} \ldots D_{s_n} Y_t \|_{2q} \leq C,$$
$$\sup_{s_1,\ldots,s_n \leq T} \| \sup_{t\leq T} |D_{s_1} \ldots D_{s_n} (X_t - Y_t) \|_{2q} \leq C h^q,$$

with C *a positive constant that depends only on* M*,* q*,* n *and* T*.*

By virtue of this Lemma, we can prove the following result about the Malliavin covariance matrix that, later on, will help us establish the convergence rate of the approximations towards the solution.

In the rest of the article we will assume that (H0), (H1) are satisfied and that either one of (H2) or (H3) is satisfied.

Lemma 3.3: *Let* ν *be a constant in* $[0,1]$ *and let* $\bar{W}$ *denote a Wiener process independent of* W*, then for any fixed* $s, t \in [0,T]$ *and* $p \in \mathbb{N}$*, we have*

$$\sup_{\nu\in[0,1]} \| \nu(Y_t - X_t) + a\bar{W}_s \|_{1,p} \leq K_1\sqrt{h} + K_2 a, \qquad \sup_{h\in(0,1]} \sup_{\nu\in[0,1]} \| \gamma^{-1}_{X_t + \nu(Y_t - X_t) + \sqrt{h}\bar{W}_s} \|_p < \infty.$$

PROOF: Let us denote by $(\bar\Omega, \bar{\mathcal F}, \bar P)$ the canonical space where $\bar W$ lives and let us define the Sobolev norms for the product space $\Omega\times\bar\Omega$ in the natural manner, that is to say (having denoted by $P' = P\times\bar P$ and $E' = E\times\bar E$)

$$\| F \|^p_{k,p} = E'(|F|^p) + \sum_{j=1}^{k}\sum_{k_1+k_2=j} E'\left[(\int_0^T \cdots \int_0^T |D^{2,k_2}_{s_j,..,s_{j-k_2}} \cdots D^{1,k_1}_{s_{j-k_2},..,s_1} F|^2 ds_j \dots ds_1)^{\frac{p}{2}}\right],$$

where $D^1 = D$ and $D^2 = \bar D$. We want to show that for any $a\in\mathbb{R}^+$,

$$\sup_{\nu\in[0,1]} \| \nu(Y_t - X_t) + a\bar W_s \|_{1,p} \le K_1\sqrt{h} + K_2 a. \tag{3.7}$$

By definition, we have

$$\begin{aligned}\| \nu(Y_t - X_t) + a\bar W_s\|^p_{1,p} =& E'(|\nu(Y_t - X_t) + a\bar W_s|^p) + E'[(\int_0^T |D_r(\nu(Y_t - X_t) + a\bar W_s)|^2 dr)^{\frac{p}{2}}] \\ &+ E'[(\int_0^T |\bar D_r(\nu(Y_t - X_t) + a\bar W_s|^2 dr)^{\frac{p}{2}}],\end{aligned}$$

which, by independence, becomes

$$\begin{aligned}\| \nu(Y_t - X_t) + a\bar W_s \|^p_{1,p} &= E'(|\nu(Y_t - X_t) + a\bar W_s|^p) \\ &+ E[(\int_0^T \nu^2 |D_r(Y_t - X_t)|^2 dr)^{\frac{p}{2}}] + \bar E[(\int_0^T a^2 |\bar D_r \bar W_s|^2 dr)^{\frac{p}{2}}] \\ &\le 2^{p-1}[\nu^p E(|Y_t - X_t|^p) + a^p \bar E(|\bar W_s|^p)] \\ &+ \nu^p E[(\int_0^T |D_r(Y_t - X_t)|^2 dr)^{\frac{p}{2}}] + (a^2 s)^{\frac{p}{2}} \\ &\le C(\nu^p \| Y_t - X_t \|^p_{1,p} + a^p \| \bar W_s \|^p_{1,p}).\end{aligned}$$

But Lemma 3.2 gives that

$$E(\sup_{t\le T} |Y_t - X_t|^p) + \sup_{r\le T} E(\sup_{t\le T} |D_r(Y_t - X_t)|^p) \le C_{1,p} h^{p/2}.$$

Thus, applying this in the previous inequality, we get

$$\| \nu(Y_t - X_t) + a\bar W_s \|^p_{1,p} \le C_{1,p} h^{\frac{p}{2}} + a^p \| \bar W_s \|^p_{1,p},$$

so our inequality is satisfied.

For the second inequality, we subdivide the proof in steps.
Step 1: By Theorem 2.2, we have already proved

$$\| \gamma^{-1}_{X_t} \|_p < \infty, \tag{3.8}$$

Step 2: we want to show

$$\| \gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s} \|_p \le \frac{1}{s}\frac{1}{a^2}. \tag{3.9}$$

Applying the definition of Malliavin covariance matrix, we have

$$\begin{aligned}\gamma_{X_t+\nu(Y_t-X_t)+a\bar{W}_s} &= \int_0^T |D_r(X_t+\nu(Y_t-X_t))|^2 dr + \int_0^T a^2|\bar{D}_r\bar{W}_s|^2 dr \\ &= \int_0^t |D_rX_t(1-\nu)+\nu D_rY_t|^2 dr + a^2 s \ge a^2 s.\end{aligned}$$

Consequently we obtain (3.9).
Step 3: Let us consider the set $A = \{|\gamma_{X_t+\nu(Y_t-X_t)+a\bar{W}_s} - \gamma_{X_t}| \le \frac{1}{2}|\gamma_{X_t}|]$. It is then clear that we have

$$\begin{aligned}E'(|\gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s}|^p) &= E'(|\gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s}|^p 1_A) + E'(|\gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s}|^p 1_{A^c}) \\ &\le 2^p E'(|\gamma^{-1}_{X_t}|^p 1_A) + E'(|\gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s}|^p 1_{A^c}) \\ &\le 2^p E'(|\gamma^{-1}_{X_t}|^p 1_A) + P(A^c)^{\frac{1}{2}} E'(|\gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s}|^{2p})^{\frac{1}{2}},\end{aligned}$$

since from (3.9) we know that $E'(|\gamma^{-1}_{X_t+\nu(Y_t-X_t)+a\bar{W}_s}|^{2p})^{\frac{1}{2}} \le \frac{1}{a^2 s}$, by taking $a=\sqrt{h}$ and using (3.8), we can conclude the proof if we notice that

$$P'(A^c) \le 2^k E'(|\gamma^{-1}_{X_t}|^k |\gamma_{X_t+\nu(Y_t-X_t)+h\bar{W}_t} - \gamma_{X_t}|^k) \le Ch^{k/2},$$

for any k. Taking k big enough, one obtains the result. $\square$

In the light of the previous lemma, we can consider the first step of our approximation procedure and obtain the following

Lemma 3.4: *With the above notation and the hypotheses of Theorem 3.1, we have*

$$\sup_{x\in\mathbb{R}} |p_t(x) - E(\phi_h(X_t-x))| \le Ch\,, \tag{3.10}$$

$$\int_{\mathbb{R}} |p_t(x) - E(\phi_h(X_t-x))| dx \le Ch, \tag{3.11}$$

with C independent of h.

Proof: In order to evaluate (3.10), as we did in Lemma 3.3 let us consider a Brownian motion $\bar{W}$, independent of the original one and let E' denote the expectation on the canonical product space, while D and $\bar{D}$ are the Malliavin derivatives with respect to W and $\bar{W}$.

The difference in (3.10) can be written as

$$p_t(x) - E(\phi_h(X_t - x)) = p_t(x) - E'(\delta_x(X_t + h^{\frac{1}{2}}\bar{W}_1)) = E'[\delta_x(X_t) - \delta_x(X_t + h^{\frac{1}{2}}\bar{W}_1)].$$

But as X_t and $X_t + h^{\frac{1}{2}}\bar{W}_1$ have smooth densities, it is known that $\phi_a(y-x) \to \delta_x(y)$ as $a \to 0$, so the last equality leads to

$$\begin{aligned} p_t(x) - E(\phi_h(X_t - x)) &= \lim_{a\to 0} E'[\phi_a(X_t - x) - \phi_a(X_t + h^{\frac{1}{2}}\bar{W}_1 - x)] \\ &= -\lim_{a\to 0} E'[\phi_a'(X_t - x)h^{\frac{1}{2}}\bar{W}_1 + \frac{1}{2}\phi_a''(\xi_t - x)h\bar{W}_1^2], \end{aligned}$$

where we used the Taylor expansion up to the second order and ξ_t represents a midpoint between X_t and $X_t + h^{\frac{1}{2}}\bar{W}_1$. By virtue of the independence between X and $\bar{W}$, we also have

$$\begin{aligned} p_t(x) - E(\phi_h(X_t - x)) &= -\lim_{a\to 0}\left[h^{\frac{1}{2}}E'(\phi_a'(X_t - x))E'(\bar{W}_1) + \frac{1}{2}E'(\phi_a''(\xi_t - x)h\bar{W}_1^2)\right] \\ &= -\lim_{a\to 0}\frac{1}{2}hE'(\phi_a''(\xi_t - x)\bar{W}_1^2). \end{aligned}$$

Given this last equality, let us focus our attention on ξ_t. We remark that for any smooth function f we may rewrite the mean value theorem for two random variables M and N as

$$f(M) - f(N) = \int_0^1 f'(M + \nu(N - M))d\nu(M - N) \tag{3.12}$$

and consequently in our case, we have

$$\begin{aligned} E'(\phi_a''(\xi_t - x)\bar{W}_1^2) &= E'(\int_0^1 \phi_a''(X_t + \nu\sqrt{h}\bar{W}_1 - x)d\nu\bar{W}_1^2) \\ &= \int_0^1 E'(\phi_a''(X_t + \nu\sqrt{h}\bar{W}_1 - x)\bar{W}_1^2)d\nu, \end{aligned} \tag{3.13}$$

where in the last passage we used Fubini's theorem. In order to prove all the statements involving a random midpoint one uses this exchange of integrals to work with specific processes rather than random midpoints.

Following [N2], for any two random variables $M, N \in \mathbb{D}^\infty$, so that $\gamma_M \in \cap_{p>1}L^p$ and $f \in \mathbb{C}_p^\infty$, the following integration by parts formula holds

$$E(f^{(m)}(M)N) = E(f(M)H_m(M,N)) \qquad \text{for} \quad m \geq 1, \tag{3.14}$$

where $H_m(M,N) = H(M, H_{m-1}(M,N))$ and

$$H_1(M,N) = H(M,N) = \delta(N\gamma_M^{-1}DM) + \bar{\delta}(N\gamma_M^{-1}\bar{D}M)$$

with δ and $\bar{\delta}$ the adjoint operators of D and $\bar{D}$ respectively.

Moreover (see [N2] page 41), for any $p > 1$, there exist indices $p_1, p_2, p_3, \alpha_1, \alpha_2$, depending on m and p and a constant $C = C(m, p, p_1, p_2, p_3)$ such that

$$\| H_m(M, N) \|_p \le C \| \gamma_M^{-1} \|_{p_1}^{\alpha_1} \| M \|_{m+1,p_2}^{\alpha_2} \| N \|_{m,p_3} \tag{3.15}$$

by $\| \cdot \|_{d,b}$ we mean the Sobolev norm, relative to $(D, \bar{D})$, as defined in section 2. Since the proof is based on Hölder's inequality for Sobolev norms, if the index $p_3 > p$ is assigned, the other two indices p_1, p_2 can be determined accordingly. We can reexpress (3.15), in a handier form, by saying that there exist integers $h, k, l, \alpha_1 \alpha_2$

$$\| H_m(M, N) \|_p \le C \| \gamma_M^{-1} \|_{lp}^{\alpha_1} \| M \|_{m+1,kp}^{\alpha_2} \| N \|_{m,hp},$$

for a properly chosen constant $C = C(m, p, h, k, l)$.

In our case, applying the integration by parts formula (3.14), we obtain

$$\begin{aligned} E'(\phi_a''(\xi_t - x)\bar{W}_1^2) &= \int_0^1 E'(\phi_a''(X_t + \nu\sqrt{h}\bar{W}_1 - x)\bar{W}_1^2) d\nu \\ &= \int_0^1 E'(\Phi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x) H_3(X_t + \nu\sqrt{h}\bar{W}_1 - x, \bar{W}_1^2)) d\nu, \end{aligned}$$

where by Φ_a we mean the Gaussian distribution function with density ϕ_a. Let us remark that by definition, H results to be independent of x, hence $H_3(X_t + \nu\sqrt{h}\bar{W}_1 - x, \bar{W}_1^2) = H_3(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2)$ for any $x \in \mathbb{R}$. Besides, as Φ is a distribution function, $0 \le \Phi_a \le 1$, so from (3.15) for some constants k, d, b, d', b' and q' we may conclude that

$$\begin{aligned} |E'(\phi_a''(\xi_t - x)\bar{W}_1^2)| &\le \int_0^1 E'(\Phi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x) \, |H_3(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2)|) d\nu \\ &\le C \int_0^1 \| \gamma_{X_t + \nu\sqrt{h}\bar{W}_1}^{-1} \|_k^q \, \| X_t + \nu\sqrt{h}\bar{W}_1 \|_{d,b}^{q'} \, \| \bar{W}_1^2 \|_{d',b'} \, d\nu. \end{aligned}$$

By the independence of $\bar{W}$ and X, for any k, $\| \gamma_{X_t + \nu\sqrt{h}\bar{W}_1}^{-1} \|_k \le \| \gamma_{X_t}^{-1} \|_k$, which is finite by Theorem 2.1. In this way this quantity is dominated independently of h and ν. Moreover $\| \bar{W}_1^2 \|_{d',b'} < \infty$ and $\| X_t + \nu\sqrt{h}\bar{W}_1 \|_{d,b} \le \| X_t \|_{d,b} + \| \nu\sqrt{h}\bar{W}_1 \|_{d,b}$ and the two terms are bounded, the first because of Lemma 3.2, the second can be made bounded independently of ν and h, if we assume without loss of generality that $h \le 1$.

Consequently we may conclude that there exists a constant C independent of h, a and x such that

$$|E'(\phi_a''(\xi_t - x)\bar{W}_1^2)| \le C,$$

that implies

$$|p_t(x) - E(\phi_h(X_t - x))| \le \frac{1}{2} C h \,,$$

which concludes the proof of (3.10)

It remains to show inequality (3.11). We have

$$
\begin{aligned}
\int_{\mathbb{R}} |p_t(x) - E(\phi_h(X_t - x))|dx &= \int_{\mathbb{R}} |\lim_{a\to 0} \frac{h}{2} E'(\phi_a''(\xi_t - x)\bar{W}_1^2)|dx \\
&= \frac{h}{2}\int_{\mathbb{R}} |\lim_{a\to 0} E'(\int_0^1 \phi_a''(X_t + \nu\sqrt{h}\bar{W}_1 - x)d\nu \bar{W}_1^2)|dx \\
&\leq \frac{h}{2}\int_{\mathbb{R}} \lim_{a\to 0} \int_0^1 |E'(\phi_a''(X_t + \nu\sqrt{h}\bar{W}_1 - x)\bar{W}_1^2)|d\nu dx \\
&= \frac{h}{2}\lim_{a\to 0} \int_{\mathbb{R}} \int_0^1 |E'(\phi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x). \\
&\qquad\qquad H_2(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2))|d\nu dx \\
&\leq \frac{h}{2}\int_0^1 E'(|H_2(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2)|)d\nu.
\end{aligned}
$$

In order to assure the interchange between the limit and the integral in the fourth passage, we are going to show that the family of functions is uniformly integrable. This will conclude the proof of (3.11).

Uniform square integrability suffices, so we want to prove that

$$
\sup_{a\in(0,1]} \int_{\mathbb{R}} \int_0^1 |E'(\phi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x)H_2(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2))|^2 d\nu dx < \infty,
$$

by exploiting the classical estimates on the exponential tails of the Gaussian density. For fixed $K \in \mathbb{R}^+$, let us divide the integral into two pieces

$$
\int_{\mathbb{R}} = \int_{\{|x|\leq K\}} + \int_{\{|x|>K\}} = I_1 + I_2.
$$

Using the same proof done for (3.10) we have that $\sup_{a\in(0,1]} I_1 < 2KC_1^2$. For I_2, let us consider $A = \{|X_t + \nu\sqrt{h}\bar{W}_1| < \frac{|x|}{2}\}$ and let us notice that if we consider the function $\Psi_a(x) = -(1 - \Phi_a(x))1_{\{x>0\}} + \Phi_a(x)1_{\{x\leq 0\}}$, then $\Psi_a'(x) = \phi_a(x)$, hence by applying the integration by parts, I_2 can be rewritten as follows and

$$
\begin{aligned}
I_2 &= \int_{\{|x|>K\}} \int_0^1 |E'(\Psi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x)H_3(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2)(1_A + 1_{A^c}))|^2 d\nu dx \\
&\leq 2\int_{\{|x|>K\}} \int_0^1 |E'(\Psi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x)H_3(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2)1_A)|^2 d\nu dx \\
&+ 2\int_{\{|x|>K\}} \int_0^1 |E'(\Psi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x)H_3(X_t + \nu\sqrt{h}\bar{W}_1, \bar{W}_1^2)1_{A^c})|^2 d\nu dx.
\end{aligned}
$$

On A we have that $|X_t + \nu\sqrt{h}\bar{W}_1 - x| > \frac{|x|}{2}$, thus for $|x|$ large enough, we can use the estimate

$$
\Psi_a(X_t + \nu\sqrt{h}\bar{W}_1 - x) \leq \exp(-\frac{x^2}{8a^2}),
$$

so that

$$\int_{\{|x|>K\}} \int_0^1 |E'(\Psi_a(X_t+\nu\sqrt{h}\bar{W}_1-x)H_3(X_t+\nu\sqrt{h}\bar{W}_1,\bar{W}_1^2))1_A)|^2 d\nu dx$$
$$\leq \int_{\{|x|>K\}} \mathrm{e}^{-\frac{x^2}{4a^2}} \int_0^1 E'(|H_3(X_t+\nu\sqrt{h}\bar{W}_1,\bar{W}_1^2))|^2) d\nu dx \leq C<\infty \qquad \forall\, a\in(0,1].$$

On A^c, it is enough to apply a Chebyshev's inequality to obtain that

$$\int_{\{|x|>K\}} \int_0^1 |E'(\Psi_a(X_t+\nu\sqrt{h}\bar{W}_1-x)H_3(X_t+\nu\sqrt{h}\bar{W}_1,\bar{W}_1^2)1_{A^c})|^2 d\nu dx$$
$$\leq \int_{\{|x|>K\}} \int_0^1 E'(|H_3(X_t+\nu\sqrt{h}\bar{W}_1,\bar{W}_1^2)|^2)P(A^c) d\nu dx$$
$$\leq \int_{\{|x|>K\}} \int_0^1 E'(|H_3(X_t+\nu\sqrt{h}\bar{W}_1,\bar{W}_1^2)|^2)\frac{2^k}{|x|^k}E(|X_t+\nu\sqrt{h}\bar{W}_1|^k) d\nu dx < C<\infty,$$

for $k>1$ and all $a\in(0,1]$. $\square$

We can pass to the second step of our procedure. This is rather more complicated than the first one and it will need several lemmas for its proof. The first one is really the key Lemma that is central to the proof of the main result. It looks a little cumbersome, but we need to state it in this generality, in order to be able to apply to both Lemmas 3.8 and 3.9, that will give the steps of the induction invoked in the proof of the main Theorem 3.7. The statement is divided into two parts, the first regarding bounded functions, while the second will give the same inequality for the particular case of the approximations $\phi_{\frac{h}{2}}$, that clearly are no longer bounded as $h\to 0$. This second part requires the introduction of yet another independent Brownian motion, to be able use the integration by parts formula.

Lemma 3.5: *Let W and $\tilde{W}$ be two independent Brownian motions, so that equations (2.1) and (3.1), defining X and Y, are driven by W, while independent copies of those, $\tilde{X}$ and $\tilde{Y}$, are driven by $\tilde{W}$. $E''=E\times\tilde{E}$ denotes the expectation on the canonical product space $\Omega\times\tilde{\Omega}$. Let V^h, Z^h be two sequences of processes adapted to the filtration generated by W, such that*

$$\begin{aligned}\sup_{s_1,\ldots,s_n\leq T} E''[\sup_{t\leq T}|D_{s_1}\ldots D_{s_n}V_t^h|^{2q}]^{1/2q} &\leq C_V\equiv C_V(T)\\ \sup_{s_1,\ldots,s_n\leq T} E''[\sup_{t\leq T}|D_{s_1}\ldots D_{s_n}Z_t^h|^{2q}]^{1/2q} &\leq C_Z\equiv C_Z(T)\end{aligned}\tag{3.16}$$

for some constants $C_V, C_Z>0$, some $q\geq 4$ and for all $n=0,1,\ldots,4$.

Moreover let

$$\alpha:\mathbb{R}^4\longrightarrow\mathbb{R},\quad \gamma:\mathbb{R}^+\times\mathbb{R}^4\longrightarrow\mathbb{R},\quad \beta:\mathbb{R}^+\times\mathbb{R}^8\longrightarrow\mathbb{R}$$

be differentiable real valued functions such that there exists positive constants C_α, C_β, C_γ, upper bounds for the following respective quantities

$$\begin{aligned} &||\alpha^{(i)}||_\infty, \text{ and } |\alpha(0,0)|, \sup_{s\in[0,T]} ||\beta_s^{(i)}||_\infty \text{ and } \sup_{s\in[0,T]} |\beta_s(0,0)| \\ &\sup_{s\in[0,T]} ||\gamma_s^{(i)}||_\infty \text{ and } \sup_{s\in[0,T]} |\gamma_s(0,0)|, \end{aligned} \tag{3.17}$$

for all $i = 1, ..., 4$ ($f^{(i)}$ denotes any partial derivative or order i).

Let us set $\underline{U}_s = (\underline{U}_s^1, \underline{U}_s^2) = ((X_{\eta(s)}, Y_{\eta(s)}, \tilde{X}_{\eta(s)}, \tilde{Y}_{\eta(s)}), (X_s, Y_s, \tilde{X}_s, \tilde{Y}_s))$, then we have

$$\left| E''\left[V_t^h \alpha(\underline{U}_t^2) \int_0^t Z_s^h \beta_s(\underline{U}_s) \int_{\eta(s)}^s \gamma_{\eta(r)}(\underline{U}_r^1) dW_r^{j_1} dW_s^{j_2} \right] \right| \le C_V C_Z C_\alpha C_\beta C_\gamma C h t, \tag{3.18}$$

where $dW_s^0 = ds$, and $(W^1, W^2) = (W, \tilde{W})$ and $j_1, j_2 = 0, 1, 2$, with C a positive constant depending only on the constant appearing in Lemma 3.2 and independent of h and all the constants $C_V, C_Z, C_\alpha, C_\beta, C_\gamma$.

Let $\bar{W}$ be a Wiener process independent of W and $\tilde{W}$ and let $E''' = E \times \tilde{E} \times \bar{E}$ denote the expectation in the cross product space supporting all 3 independent processes. Let us take $\alpha(\underline{U}_t^2) = \alpha(X_t, Y_t) = \phi_{\frac{h}{2}}^{(p)}(X_t + \nu(Y_t - X_t) + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)$, for $p \in \{0, 1\}$. Then assuming that (3.16) holds with $q \ge 32$ and that β and γ verify (3.17) for $i = 1, ..., p+3$, we have that

$$\left| E''' \left[V_t^h \alpha(\underline{U}_t^2) \int_0^t Z_s^h \beta_s(\underline{U}_s) \int_{\eta(s)}^s \gamma_{\eta(r)}(\underline{U}_r^1) dW_r^{j_1} dW_s^{j_2} \right] \right| \le C_V C_Z C_\beta C_\gamma C h t, \tag{3.19}$$

uniformly for $\nu \in [0, 1]$ and .

The constant C in (3.19) is different from the one in (3.18) and in the second part we are clearly no longer assuming α bounded. From now on, we will use the notation $Z_t^{\nu,\bar{W}} = X_t + \nu(Y_t - X_t) + \sqrt{h}\bar{W}_{\frac{1}{2}}$.

PROOF: We will prove (3.18) only when $j_1 = 1, j_2 = 1$, which is computationally the most cumbersome case, all the others can be treated similarly by applying the integration by parts once or twice less. Indeed we have to use integration by parts once less each time we have a time integral. If instead we are considering double integrals with respect to the two independent Brownian motions, the procedure we are going to describe simplifies, as some terms will become zero, because the processes V^h and Z^h depend only upon W.

To simplify notation, we are going to omit the arguments of the functions.

Applying the integration by parts formula of Malliavin Calculus with respect to W, we have

$$|E''[V_t^h\alpha\int_0^t Z_s^h\beta_s\int_{\eta(s)}^s\gamma_{\eta(r)}dW_rdW_s]| = |E''[\int_0^t D_s\{V_t^h\alpha\}Z_s^h\beta_s\int_{\eta(s)}^s\gamma_{\eta(r)}dW_rds]|$$

$$= |E''[\int_0^t\int_{\eta(s)}^s D_r[D_s\{V_t^h\alpha\}Z_s^h\beta_s]\gamma_{\eta(r)}drds]|$$

$$\leq \int_0^t\int_{\eta(s)}^s |E''[D_r[D_s\{V_t^h\alpha\}Z_s^h\beta_s]\gamma_{\eta(r)}]|drds\,.$$

It is then clear that to obtain (3.18), it suffices to show that

$$\sup_{\substack{s\in[0,t]\\ r\in(\eta(s),s]}} |E''[D_r[D_s\{V_t^h\alpha\}Z_s^h\beta_s]\gamma_{\eta(r)}]| \leq C_VC_ZC_\alpha C_\beta C_\gamma C,$$

where C is a positive constant that depends only on T and the constant appearing in Lemma 3.2.

Applying assumption (3.16) and Hölder's inequality, we get

$$\begin{aligned}
&|E''(\,D_r\{\,D_s\{V_t^h\alpha\}Z_s^h\beta_s\,\}\gamma_{\eta(r)})| \\
\leq& E''\left[|D_rD_s\{V_t^h\alpha\}Z_s^h\beta_s\gamma_{\eta(r)}| + |D_s\{V_t^h\alpha\}D_rZ_s^h\beta_s\gamma_{\eta(r)}| + |D_s\{V_t^h\alpha\}Z_s^hD_r\beta_s\gamma_{\eta(r)}|\right] \\
\leq& \|\gamma_{\eta(r)}\|_4\left\{\|Z_s^h\|_4(\|\beta_s\|_4\|D_rD_s\{V_t^h\alpha\}\|_4 + \|D_r\beta_s\|_4\|D_s\{V_t^h\alpha\}\|_4)\right. \\
&\left.+\|D_rZ_s^h\|_4\|\beta_s\|_4\|D_s\{V_t^h\alpha\}\|_4\right\}.
\end{aligned}$$

From now on, we will denote each component of $\underline{U}$ by U^i for $i = 1, ..., 8$. We are going to analyze each single term, indeed by assumptions (3.16) and (3.17) and Hölder's inequality, we may dominate each of them in the following manner

(A) $$\|Z_s^h\|_4 \leq C_Z, \quad \|D_rZ_s^h\|_4 \leq C_Z;$$

(B) $$||\gamma_{\eta(r)}||_4 \leq ||\sum_{i=1}^4\frac{\partial\gamma_{\eta(r)}}{\partial x_i}U_r^i||_4 + ||\gamma_{\eta(r)}(\underline{0})||_4 \leq C_\gamma(\sum_{i=1}^4||U_r^i||_4 + 1);$$

(C) $$\|\beta_s\|_4 \leq ||\sum_{i=1}^8\frac{\partial\beta_s}{\partial x_i}U_s^i||_4 + \|\beta_s(\underline{0})\|_4 \leq C_\beta(\sum_{i=1}^8||U_s^i||_4 + 1);$$

(D) $$\|D_r\beta_s\|_4 \leq ||\frac{\partial\beta_s}{\partial x_5}D_rU_s^5 + \frac{\partial\beta_s}{\partial x_6}D_rU_s^6||_8 \leq C_\beta(||D_r\tilde{X}_s||_4 + ||D_r\tilde{Y}_s||_4);$$

(E) $$\begin{aligned}
||D_s(V_t^h\alpha)||_4 &\leq ||D_sV_t^h(\sum_{i=1}^4\frac{\partial\alpha}{\partial x_i}U_t^{i+4} + \alpha(\underline{0}))||_4 + ||V_t^h(\frac{\partial\alpha}{\partial x_1}D_sX_t + \frac{\partial\alpha}{\partial x_2}D_sY_t)||_4 \\
&\leq C_\alpha\{||D_sV_t^h||_8(1+\sum_{i=5}^8||U_t^i||_8) + ||V_t^h||_8(||D_sX_t||_8 + ||D_sY_t||_8)\} \\
&\leq C_VC_\alpha[\sum_{i=5}^8||U_t^i||_8 + 1 + ||D_sX_t||_8 + ||D_sY_t||_8];
\end{aligned}$$

$$
\text{(F)}\qquad
\begin{aligned}
||D_rD_s\{V_t^h\alpha\}||_4 \leq{}& ||D_rD_sV_t^h||_8(\sum_{i=1}^{4}||\frac{\partial\alpha}{\partial x_i}U_t^{i+4}||_8+||\alpha(\underline{0})||_8)\\
&+||D_sV_t^h||_8||\sum_{i=1}^{2}\frac{\partial\alpha}{\partial x_i}D_rU_t^{i+4}||_8+||D_rV_t^h||_8||\sum_{i=1}^{2}\frac{\partial\alpha}{\partial x_i}D_sU_t^{i+4}||_8\\
&+||V_t^h||_8[||\sum_{i,j=1,2}\frac{\partial^2\alpha}{\partial x_i\partial x_j}D_sU_t^{i+4}D_rU_t^{j+4}||_8+||\sum_{i,j=1,2}\frac{\partial\alpha}{\partial x_i}D_rD_sU_t^{i+4}||_8]\\
\leq{}& C_VC_\alpha\{\sum_{i=5}^{8}||U_t^i||_8+1+2(||D_sU_t^5||_8+||D_sU_t^6||_8)\\
&+(||D_sU_t^5||_{16}+||D_sU_t^6||_{16})^2+||D_rD_sU_t^5||_8+||D_rD_sU_t^6||_8\}.
\end{aligned}
$$

By virtue of all the previous estimates and using Lemma 3.2, we may conclude that X and Y together with their Malliavin derivatives are bounded in the L^p norms ($p \leq 8$) uniformly in t, let us say by a common constant C, so we finally get

$$
\sup_{\substack{s\in[0,t]\\ r\in[\eta(s),s]}}|E''[D_r[D_s\{V_t^h\alpha\}Z_s^h\beta_s]\gamma_{\eta(r)}]|\leq C_VC_ZC_\alpha C_\beta C_\gamma(4C+1)(8C+1)(4C^2+20C+2).
$$

To prove the second result in the statement, with $\phi_{\frac{h}{2}}^{(p)}$ in place of α, we restrict to the case $j_1=2, j_2=1$ (also to give an idea on how to deal with a different case) and we denote by $Z_t^{\nu,\bar{W}}=X_t+\nu(Y_t-X_t)+\sqrt{h}\bar{W}_{\frac{1}{2}}$. The main difference with the previous proof lies on the fact that we lose the uniform bounds on the derivatives of α, but a double application of integration by parts will help us. Again by integration by parts, the problem is reduced to showing that $|E'''[\tilde{D}_r\{D_s[V_t^h\phi_{\frac{h}{2}}^{(p)}(Z_t^{\nu,\bar{W}}-x)]Z_s^h\beta_s\}\gamma_{\eta(r)}]|$ is bounded uniformly in s,r and ν. Carrying the calculations out, we get

$$
\begin{aligned}
&|E'''[\tilde{D}_r\{D_s[V_t^h\phi_{\frac{h}{2}}^{(p)}(Z_t^{\nu,\bar{W}}-x)]Z_s^h\beta_s\}\gamma_{\eta(r)}]|\\
\leq{}&|E'''[\phi_{\frac{h}{2}}^{(p)}(Z_t^{\nu,\bar{W}}-x)\gamma_{\eta(r)}D_sV_t^hZ_s^h\tilde{D}_r\beta_s]|+|E'''[\phi_{\frac{h}{2}}^{(p+1)}(Z_t^{\nu,\bar{W}}-x)D_sZ_t^{\nu,\bar{W}}\gamma_{\eta(r)}V_t^hZ_s^h\tilde{D}_r\beta_s]|\\
={}&|E'''[\Phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)H_{p+1}(Z_t^{\nu,\bar{W}},N^1)]|+|E'''[\Phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)H_{p+2}(Z_t^{\nu,\bar{W}},N^2)]|,
\end{aligned}
$$

where N^1 and N^2 have been obviously defined.

By applying (3.15) to the above terms we may conclude

$$
\begin{aligned}
&|E'''[\tilde{D}_r\{D_s[V_t^h\phi_{\frac{h}{2}}^{(p)}(Z_t^{\nu,\bar{W}}-x)]Z_s^h\beta_s\}\gamma_{\eta(r)}]|\\
\leq{}& C_{p+1}\|\gamma^{-1}_{Z_t^{\nu,\bar{W}}}\|_{q_1}^{m_1}\|Z_t^{\nu,\bar{W}}\|_{p+2,q_2}^{m_2}\|N^1\|_{p+1,q_3}+C_{p+2}\|\gamma^{-1}_{Z_t^{\nu,\bar{W}}}\|_{d_1}^{n_1}\|Z_t^{\nu,\bar{W}}\|_{p+3,d_2}^{n_2}\|N^2\|_{p+2,d_3}
\end{aligned}
$$

but $\|\gamma^{-1}_{Z_t^{\nu,\bar{W}}}\|_{q_1}$, $\|\gamma^{-1}_{Z_t^{\nu,\bar{W}}}\|_{d_1}$, are bounded by virtue of Lemma 3.3, moreover we know $\|Z_t^{\nu,\bar{W}}\|_{p+3,q_2}\leq||X_t||_{p+3,q_2}+||Y_t||_{p+3,q_2}<+\infty$ and by the increasingness of the Sobolev norms, this implies that also the term $\|Z_t^{\nu,\bar{W}}\|_{p+2,d_2}$ is bounded.

So it remains to evaluate $\| N^1 \|_{p+1,q_3}$ and $\| N^2 \|_{p+2,d_3}$, we will show the boundedness only of the first term, as the proof is the same for both.

If we apply the Hölder's inequality for Sobolev norms, we obtain

$$\begin{aligned} ||N^1||_{p+1,q_3} &\leq ||\gamma_{\eta(r)} D_s V_t^h Z_s^h \tilde{D}_r \beta_s||_{p+1,q_3} \\ &\leq ||\gamma_{\eta(r)}||_{p+1,b_1} ||D_s V_t^h||_{p+1,b_2} ||Z_s^h||_{p+1,b_3} ||\tilde{D}_r \beta_s||_{p+1,b_4}\} \\ &\leq C_V C_Z ||\gamma_{\eta(r)}||_{p+1,b_1} ||\tilde{D}_r \beta_s||_{p+1,b_4}\}, \end{aligned}$$

where $\frac{1}{b_1} + \frac{1}{b_2} + \frac{1}{b_3} + \frac{1}{b_4} = \frac{1}{q_3}$. On the other hand it is easy to prove that, if f is a smooth function with its derivatives and $|f(0)|$ uniformly bounded by a constant A and G is a random variable, then

$$||f(G)||_{p+1,q} \leq \Lambda A ||G||_{p+1,nq},$$

for appropriate Λ and n. Consequently, in our case we have

$$\begin{aligned} ||\gamma_{\eta(r)}||_{p+1,b_1} &\leq \Lambda_1 C_\gamma (||X_{\eta(r)}||_{p+1,nb_1} + ||Y_{\eta(r)}||_{p+1,nb_1}) \\ ||\tilde{D}_r \beta_s||_{p+1,b_4} &\leq \Lambda_2 C_\beta \rho(C), \end{aligned}$$

for some fixed polynomial function ρ, constants Λ_1, Λ_2 and integers m, n, which concludes the proof. $\square$

Remark 3.6:

(a) The same technique applies also to prove

$$|E'(\phi^{(p)}_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x) V_t^h \int_0^t Z_s^h \beta_s dW_s^j)| \leq \rho(C) C_V C_\beta \int_0^t C_Z(s) ds, \qquad j = 0,1 \quad p = 0,1$$

for some properly chosen polynomial ρ, when β depends only on X, Y and verifies (3.17).

Indeed, for j=1

$$|E'(\phi^{(p)}_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x) V_t^h \int_0^t Z_s^h \beta_s dW_s^j)| \leq \int_0^t \Big[||\gamma^{-1}_{Z_t^{\nu,\bar{W}}}||^{m_1}_{q_1} ||Z_t^{\nu,\bar{W}}||^{m_2}_{p+2,q_2} ||D_s V_t^h Z_s^h \beta_s||_{p+1,q_3}$$

$$+ ||\gamma^{-1}_{Z_t^{\nu,\bar{W}}}||^{n_1}_{d_1} ||Z_t^{\nu,\bar{W}}||^{n_2}_{p+3,d_2} ||D_s Z_t^{\nu,\bar{W}} V_t^h Z_s^h \beta_s||_{p+2,d_3} \Big] ds$$

and we may proceed as before.

(b) Another important fact we would like to point out is that the previous Lemma holds even when α is a smooth random function independent of $W, \tilde{W}$, with the bounds in (3.17) holding for every ω.

In this case, the same proof goes through, since the Malliavin derivatives of α do not intervene in the computation, but only the spatial ones. Therefore the right handside of (3.18) will result a random function bounded by the same constant uniformly in ω.

(c) Another point that we would like to remark is that in the previous proof one might assume a lower degree of integrability in (3.16), provided one chooses to penalize more the other terms, when applying Hölder's inequality.

The main result for the second step is summarized in the following

Theorem 3.7: *Under the same hypotheses as in Theorem 3.1, the following inequality holds*

$$\sup_{x\in\mathbb{R}} |E(\phi_h(X_t - x) - \phi_h(Y_t - x))| \leq Ch\,, \tag{3.20}$$

with C independent of h.

Proof: By applying the mean value theorem to the difference under expectation in (3.20), we have

$$\begin{aligned} E(\phi_h(X_t - x) - \phi_h(Y_t - x)) &= E'(\phi_{\frac{h}{2}}(X_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x) - \phi_{\frac{h}{2}}(Y_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)) \\ &= E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)(X_t - Y_t)), \end{aligned}$$

where ξ_t^1 is a random midpoint between X_t and Y_t. From the definitions of X and Y, it is easy to see that the difference verifies the following equation

$$\begin{aligned} X_t - Y_t = &\int_0^t [a(X_s, F(X_s; u_s)) - a(Y_{\eta(s)}, F(Y_{\eta(s)}; v_{\eta(s)}))]ds \\ &+ \int_0^t [b(X_s, G(X_s; u_s)) - b(Y_{\eta(s)}, G(Y_{\eta(s)}; v_{\eta(s)}))]dW_s\,. \end{aligned}$$

Let us remark that, due to the regularity of the kernels f and g, the coefficients F, G result differentiable, hence adding and subtracting the proper terms in the above expression and applying the mean value theorem on each of the differences we obtain

$$\begin{aligned} X_t - Y_t = &\int_0^t \{a_x(\xi_s^2, F(X_s; u_s))(X_s - Y_s) + a_y(Y_s, \eta_s^1)[F(X_s; u_s) - F(Y_s; v_s)]\}ds \\ &+ \int_0^t \{a_x(\zeta_s^1, F(Y_s; v_s))(Y_s - Y_{\eta(s)}) + a_y(Y_{\eta(s)}, \theta_s^1)[F(Y_s; v_s) - F(Y_{\eta(s)}; v_{\eta(s)})]\}ds \\ &+ \int_0^t \{b_x(\xi_s^3, G(X_s; u_s))(X_s - Y_s) + b_y(Y_s, \eta_s^2)[G(X_s; u_s) - G(Y_s; v_s)]\}dW_s \\ &+ \int_0^t \{b_x(\zeta_s^2, G(Y_s; v_s))(Y_s - Y_{\eta(s)}) + b_y(Y_{\eta(s)}, \theta_s^2)[G(Y_s; v_s) - G(Y_{\eta(s)}; v_{\eta(s)})]\}dW_s. \end{aligned}$$

From now on we adopt $[V; Z]$ as standard notation to indicate the interval with the random variables Z, V as endpoints, therefore in the above we have that

$$\begin{aligned} &\xi_s^2,\ \xi_s^3\ \in\ [X_s; Y_s]\,, \qquad \zeta_s^1,\ \zeta_s^2 \in\ [Y_s; Y_{\eta(s)}], \\ &\eta_s^1\ \in\ [F(X_s; u_s); F(Y_s; v_s)]\,, \quad \eta_s^2\ \in\ [G(X_s; u_s); G(Y_s; v_s)] \\ &\theta_s^1\ \in\ [F(Y_s; v_s); F(Y_{\eta(s)}, v_{\eta(s)})]\,, \quad \theta_s^2\ \in\ [G(Y_s; v_s); G(Y_{\eta(s)}, v_{\eta(s)})], \end{aligned}$$

where the midpoints are really to be intended in the notation of formula (3.12). By adding and subtracting $F(Y_s;u_s)$ in the second term of the first time integral, $G(Y_s;u_s)$ in the second term of the first Brownian integral and applying once again the mean value theorem to those, we get

$$\begin{aligned}
X_t - Y_t = &\int_0^t [a_x(\xi_s^2, F(X_s;u_s)) + a_y(Y_s,\eta_s^1)F'(\xi_s^4;u_s)](X_s - Y_s)ds \\
&+ \int_0^t \{a_y(Y_s,\eta_s^1)[F(Y_s;u_s) - F(Y_s;v_s)]\}ds \\
&+ \int_0^t \{a_x(\zeta_s^1, F(Y_s;v_s))(Y_s - Y_{\eta(s)}) + a_y(Y_{\eta(s)},\theta_s^1)[F(Y_s;v_s) - F(Y_{\eta(s)};v_{\eta(s)})]\}ds \\
&+ \int_0^t [b_x(\xi_s^3, G(X_s;u_s)) + b_y(Y_s,\eta_s^2)G'(\xi_s^5;u_s)](X_s - Y_s)dW_s \\
&+ \int_0^t \{b_y(Y_s,\eta_s^2)[G(Y_s;u_s) - G(Y_s;v_s)]\}dW_s \\
&+ \int_0^t \{b_x(\zeta_s^2, G(Y_s;v_s))(Y_s - Y_{\eta(s)}) + b_y(Y_{\eta(s)},\theta_s^2)[G(Y_s;v_s) - G(Y_{\eta(s)};v_{\eta(s)})]\}dW_s,
\end{aligned}$$

with $\xi_s^4,\ \xi_s^5 \in [X_s;Y_s]$. For simplicity of notation, from now on we set

$$\begin{aligned}
\alpha_s &= [a_x(\xi_s^2, F(X_s;u_s)) + a_y(Y_s,\eta_s^1)F'(\xi_s^4;u_s)], \\
\beta_s &= [b_x(\xi_s^3, G(X_s;u_s)) + b_y(Y_s,\eta_s^2)G'(\xi_s^5;u_s)], \\
H_t &= \int_0^t \{a_y(Y_s,\eta_s^1)[F(Y_s;u_s) - F(Y_s;v_s)] + a_y(Y_{\eta(s)},\theta_s^1)[F(Y_s;v_s) - F(Y_{\eta(s)};v_{\eta(s)})]\}ds \\
&\quad + \int_0^t \{b_y(Y_s,\eta_s^2)[G(Y_s;u_s) - G(Y_s;v_s)] + b_y(Y_{\eta(s)},\theta_s^2)[G(Y_s;v_s) - G(Y_{\eta(s)};v_{\eta(s)})]\}dW_s, \\
dK_s &= a_x(\zeta_s^1, F(Y_s;v_s))ds + b_x(\zeta_s^2, G(Y_s;v_s))dW_s\ ;\ K_0 = 0.
\end{aligned}$$

With this new notation, the above equation becomes

$$X_t - Y_t = \int_0^t (X_s - Y_s)(\alpha_s ds + \beta_s dW_s) + H_t + \int_0^t (Y_s - Y_{\eta(s)})dK_s,$$

whose explicit solution is given by

$$X_t - Y_t = \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\{dH_s + (Y_s - Y_{\eta(s)})dK_s - d[\int_0^\cdot \beta_s dW_s, H + (Y - Y_\eta)\cdot K]_s\}, \tag{3.21}$$

where $\mathcal{E}_t$ denotes $e^{\int_0^t(\alpha_s - \beta_s^2/2)ds + \int_0^t \beta_s dW_s}$. In order to simplify even further and to regroup the terms in ds and in dW_s, we consider the process $U_t = \mathcal{E}_t^{-1}(X_t - Y_t)$. With a few computations, from the definition of H, (3.21) can be rewritten as

$$\begin{aligned}
U_t = & \int_0^t \mathcal{E}_s^{-1}(Y_s - Y_{\eta(s)})[dK_s - \beta_s b_x(\zeta_s^2, G(Y_s; v_s))ds] \\
& + \int_0^t \mathcal{E}_s^{-1}\{a_y(Y_s, \eta_s^1)[F(Y_s; u_s) - F(Y_s; v_s)] + a_y(Y_{\eta(s)}, \theta_s^1)[F(Y_s; v_s) - F(Y_{\eta(s)}; v_{\eta(s)})] \\
& \quad - \beta_s(b_y(Y_s, \eta_s^2)[G(Y_s; u_s) - G(Y_s; v_s)] + b_y(Y_{\eta(s)}, \theta_s^2)[G(Y_s; v_s) - G(Y_{\eta(s)}; v_{\eta(s)})])\}ds \\
& + \int_0^t \mathcal{E}_s^{-1}\{b_y(Y_s, \eta_s^2)[G(Y_s; u_s) - G(Y_s; v_s)] + b_y(Y_{\eta(s)}, \theta_s^2)[G(Y_s; v_s) - G(Y_{\eta(s)}; v_{\eta(s)})]\}dW_s,
\end{aligned}$$

On the other hand, the differences in F and G can be reformulated making use of their respective kernels. Indeed if we introduce independent copies of X and Y, say $\tilde{X}$ and $\tilde{Y}$ and the canonical space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ where they live, we can look at those differences in the following manner

$$\begin{aligned}
F(Y_s; u_s) - F(Y_s; v_s) &= \tilde{E}(f(Y_s, \tilde{X}_s)) - \tilde{E}(f(Y_s, \tilde{Y}_s)) = \tilde{E}(f_y(Y_s, \tilde{\xi}_s^1)(\tilde{X}_s - \tilde{Y}_s)) \\
G(Y_s; u_s) - G(Y_s; v_s) &= \tilde{E}(g_y(Y_s, \tilde{\xi}_s^2)(\tilde{X}_s - \tilde{Y}_s)) \\
F(Y_s; v_s) - F(Y_{\eta(s)}; v_{\eta(s)}) &= \tilde{E}(f(Y_s, \tilde{Y}_s) - f(Y_{\eta(s)}, \tilde{Y}_{\eta(s)})) \\
&= \tilde{E}(f(Y_s, \tilde{Y}_s) - f(Y_s, \tilde{Y}_{\eta(s)}) + f(Y_s, \tilde{Y}_{\eta(s)}) - f(Y_{\eta(s)}, \tilde{Y}_{\eta(s)})) \\
&= \tilde{E}(f_x(\zeta_s^3, \tilde{Y}_{\eta(s)})(Y_s - Y_{\eta(s)}) + f_y(Y_s, \tilde{\zeta}_s^1)(Y_s - Y_{\eta(s)}) \\
G(Y_s; v_s) - G(Y_{\eta(s)}; v_{\eta(s)}) &= \tilde{E}(g(Y_s, \tilde{Y}_s) - g(Y_{\eta(s)}, \tilde{Y}_{\eta(s)})) \\
&= \tilde{E}(g_x(\zeta_s^4, \tilde{Y}_{\eta(s)})(Y_s - Y_{\eta(s)}) + g_y(Y_s, \tilde{\zeta}_s^2)(Y_s - Y_{\eta(s)})),
\end{aligned}$$

where $\tilde{E}$ denotes the expectation in $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and where we used once again the mean value theorem, with $\tilde{\xi}_s^1, \tilde{\xi}_s^2 \in [\tilde{X}_s; \tilde{Y}_s]$ and $\zeta_s^3,\ \zeta_s^4 \in [Y_{\eta(s)}; Y_s]$ and $\tilde{\zeta}_s^1,\ \tilde{\zeta}_s^2 \in [\tilde{Y}_{\eta(s)}; \tilde{Y}_s]$. Similarly, if we take an independent copy of $\mathcal{E}$, say $\tilde{\mathcal{E}}$, the above equation for U_t is transformed into

$$\begin{aligned}
U_t = & \int_0^t \mathcal{E}_s^{-1}(Y_s - Y_{\eta(s)})[dK_s - \beta_s b_x(\zeta_s^2, G(Y_s; v_s))ds] \\
& + \int_0^t \mathcal{E}_s^{-1}[a_y(Y_s, \eta_s^1)\tilde{E}(f_y(Y_s, \tilde{\xi}_s^1)\tilde{\mathcal{E}}_s\tilde{U}_s) - \beta_s b_y(Y_s, \eta_s^2)\tilde{E}(g_y(Y_s, \tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_s)]ds \\
& + \int_0^t \mathcal{E}_s^{-1}(Y_s - Y_{\eta(s)})[a_y(Y_{\eta(s)}, \theta_s^1)\tilde{E}(f_x(\zeta_s^3, \tilde{Y}_{\eta(s)})) - \beta_s b_y(Y_{\eta(s)}, \theta_s^2)\tilde{E}(g_x(\zeta_s^4, \tilde{Y}_{\eta(s)}))]ds \\
& + \int_0^t \mathcal{E}_s^{-1}[a_y(Y_{\eta(s)}, \theta_s^1)\tilde{E}(f_y(Y_s, \tilde{\zeta}_s^1)(\tilde{Y}_{\eta(s)} - \tilde{Y}_s)) - \beta_s b_y(Y_{\eta(s)}, \theta_s^2)\tilde{E}(g_y(Y_s, \tilde{\zeta}_s^2)(\tilde{Y}_{\eta(s)} - \tilde{Y}_s))]ds \\
& + \int_0^t \mathcal{E}_s^{-1} b_y(Y_s, \eta_s^2)\tilde{E}(g_y(Y_s, \tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_s)dW_s \\
& + \int_0^t \mathcal{E}_s^{-1} b_y(Y_{\eta(s)}, \theta_s^2)[(Y_s - Y_{\eta(s)})\tilde{E}(g_x(\zeta_s^4, \tilde{Y}_{\eta(s)})) + \tilde{E}(g_y(Y_s, \tilde{\zeta}_s^2)(\tilde{Y}_{\eta(s)} - \tilde{Y}_s))]dW_s.
\end{aligned}$$

We are finally in condition to rearrange the terms and obtain a simpler form for (3.21)

$$
\begin{aligned}
U_t = & \int_0^t \mathcal{E}_s^{-1}[a_y(Y_s,\eta_s^1)\tilde{E}(f_y(Y_s,\tilde{\xi}_s^1)\tilde{\mathcal{E}}_s\tilde{U}_s) - \beta_s b_y(Y_s,\eta_s^2)\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_s)]ds \\
& + \int_0^t \mathcal{E}_s^{-1} b_y(Y_s,\eta_s^2)\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_s)dW_s + \int_0^t \mathcal{E}_s^{-1}dZ_s\,,
\end{aligned}
\tag{3.22}
$$

where we set

$$
\begin{aligned}
dZ_s &= (Y_s - Y_{\eta(s)})(A_s ds + B_s dW_s) + \tilde{E}((\tilde{Y}_s - \tilde{Y}_{\eta(s)})\tilde{A}_s)ds + \tilde{E}((\tilde{Y}_s - \tilde{Y}_{\eta(s)})\tilde{B}_s)dW_s \\
B_s &= b_x(\zeta_s^2, G(Y_s;v_s)) + b_y(Y_{\eta(s)},\theta_s^2)\tilde{E}(g_x(\zeta_s^4,\tilde{Y}_{\eta(s)})) \\
A_s &= a_x(\zeta_s^1, F(Y_s;v_s)) + a_y(Y_{\eta(s)},\theta_s^1)\tilde{E}(f_x(\zeta_s^3,\tilde{Y}_{\eta(s)})) - \beta_s B_s \\
\tilde{B}_s &= b_y(Y_{\eta(s)},\theta_s^2)g_y(Y_s,\tilde{\zeta}_s^2) \\
\tilde{A}_s &= a_y(Y_{\eta(s)},\theta_s^1)f_y(Y_s,\tilde{\zeta}_s^1) - \beta_s\tilde{B}_s\,.
\end{aligned}
$$

It is easy to show that equation (3.22) has a unique solution and that the sequence of iterates defined as

$$
\begin{aligned}
U_k(t) = & \int_0^t \mathcal{E}_s^{-1}\tilde{E}\left([a_y(Y_s,\eta_s^1)f_y(Y_s,\tilde{\xi}_s^1) - \beta_s b_y(Y_s,\eta_s^2)g_y(Y_s,\tilde{\xi}_s^2)]\tilde{\mathcal{E}}_s\tilde{U}_{k-1}(s)\right)ds \\
& + \int_0^t \mathcal{E}_s^{-1}b_y(Y_s,\eta_s^2)\tilde{E}\left(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_{k-1}(s)\right)dW_s + U_0(t) \\
U_0(t) = & \int_0^t \mathcal{E}_s^{-1}dZ_s
\end{aligned}
\tag{3.23}
$$

converges to the solution (see [KO]).

By virtue of our initial remark, we can say that our proof is complete if we show that there exists a constant R independent of k, x and t such that

$$
|E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_k(t))| \le h\sum_{j=1}^k \frac{(Rt)^j}{j!}. \tag{3.24}
$$

Then by dominated convergence theorem, this implies that

$$
|E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U(t))| = \lim_{k\to\infty}|E'(\phi'_{\frac{h}{2}}(\xi_t^1 - x)\mathcal{E}_t U_k(t))| \le h e^{RT}
$$

and the first part of the theorem is proven. $\square$

In order to prove (3.24), we proceed by induction, the first step being carried out in the next lemma and the general case in Lemma 3.9.

Lemma 3.8: *Let ξ_t^1 and $U_0(t) = \int_0^t \mathcal{E}_s^{-1} dZ_s$ be both defined as before. Then there exists a deterministic constant A depending on M, but independent of t, x, U_0, such that the following holds*

$$|E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_0(t))| \leq Ath, \qquad |\tilde{E}(u(\tilde{X}_t, \tilde{Y}_t)\tilde{\mathcal{E}}_t \tilde{U}_0(t))| \leq Ath. \tag{3.25}$$

Here $u : \Omega \times \mathbb{R}^2 \to \mathbb{R}$ is any smooth random measurable function with its first 4 derivatives bounded by M uniformly in Ω, independent of $\tilde{W}$.

PROOF: Recalling the definition of Z, we can rewrite

$$\begin{aligned}
&|E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_0(t))| \\
\leq &|E'\left[\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\{(Y_s - Y_{\eta(s)})A_s + \tilde{E}((\tilde{Y}_s - \tilde{Y}_{\eta(s)})\tilde{A}_s)\}ds\right]| \\
+ &|E'\left[\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\{(Y_s - Y_{\eta(s)})B_s + \tilde{E}((\tilde{Y}_s - \tilde{Y}_{\eta(s)})\tilde{B}_s)\}dW_s\right]|.
\end{aligned}$$

The four terms in the above expression are qualitatively very similar (due to the hypotheses on the coefficients) and they differ basically only in the integrators. We show inequality (3.25) only for the fourth one, this is the most complicated term since it contains both Brownian motions. The proof of all the other terms runs along similar lines. As an independent copy of Y, $\tilde{Y}$ must verify an analogous equation

$$\tilde{Y}_s - \tilde{Y}_{\eta(s)} = a(\tilde{Y}_{\eta(s)}, F(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))(s - \eta(s)) + b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))(\tilde{W}_s - \tilde{W}_{\eta(s)}),$$

so substituting the latter in the fourth term of the previous inequality, this becomes

$$\begin{aligned}
&E'\left(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\tilde{E}\left[\tilde{B}_s a(\tilde{Y}_{\eta(s)}, F(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))\int_{\eta(s)}^s dr\right] dW_s\right) \\
+&E'\left(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\tilde{E}\left[\tilde{B}_s b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))\int_{\eta(s)}^s d\tilde{W}_r\right] dW_s\right).
\end{aligned}$$

Again we look only at the last term, since the other can be treated similarly. As we already mentioned, the midpoint ξ_t^1 is to be understood in the sense of expression (3.12), so recalling the definition of $Z_t^{\nu,\bar{W}}$, under the expectation E''' on $\Omega \times \bar{\Omega} \times \tilde{\Omega}$ (therefore $E''' = E' \times \tilde{E}$), we have

$$\begin{aligned}
&E'\left(\int_0^1 \phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)d\nu\, \mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\tilde{E}\left[\tilde{B}_s b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))\int_{\eta(s)}^s d\tilde{W}_r\right] dW_s\right) \\
= &\int_0^1 E'''\left(\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\tilde{B}_s b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))\int_{\eta(s)}^s d\tilde{W}_r dW_s\right) d\nu,
\end{aligned} \tag{3.26}$$

and we are in condition to apply Lemma 3.5. If we recall the definition of $\tilde{B}$ and we translate the midpoints θ_s^2 and $\tilde{\zeta}_s^2$ there appearing in the notation (3.12), we obtain that this last term can be actually expressed as

$$\int_0^1 E'''(\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}- x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b_y(Y_{\eta(s)}, \theta_s^2) g_y(Y_s, \tilde{\zeta}_s^2) \int_{\eta(s)}^s b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))d\tilde{W}_r dW_s)d\nu$$

$$= \int_0^1 \int_0^1 \int_0^1 E'''[\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}- x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b_y(Y_{\eta(s)}, (1-\epsilon)G(Y_{\eta(s)}; v_{\eta(s)}) + \epsilon G(Y_s; v_s))$$

$$\cdot\, g_y(Y_s, (1-\mu)\tilde{Y}_{\eta(s)} + \mu \tilde{Y}_s) \int_{\eta(s)}^s b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)}))d\tilde{W}_r dW_s] d\mu d\epsilon d\nu.$$

Indeed, by virtue of hypothesis (H0), the functions

$$\gamma(x_1) = b(x_1, \int_{\mathbb{R}} g(x_1, z) dv_{\eta(s)})$$

$$\beta(x_1, x_2, x_3, x_4) = b_y(x_1, (1-\epsilon)\int_{\mathbb{R}} g(x_1, z)dv_{\eta(s)} + \epsilon\int_{\mathbb{R}} g(x_1, z)dv_s) g_y(x_3, (1-\mu)x_2 + \mu x_4)$$

respectively applied to $Y_{\eta(s)}$ and $(Y_{\eta(s)}, \tilde{Y}_{\eta(s)}, Y_s, \tilde{Y}_s)$, verify condition (3.17). They are differentiable four times and the derivative of order i of each of them is bounded by $C_2 = 2^{2(i+1)}M^{2(i+2)}$, in the worst of cases. Besides $\mathcal{E}_t$ and its inverse are solutions to SDE's with smooth initial condition and coefficients with bounded spatial derivatives. Therefore it is not difficult to prove that they satisfy for $n = 0, 1, \ldots, 4$ and $q \in \mathbb{N}$ (see [N1], theorem 2.2.2),

$$\sup_{s_1,\ldots,s_n \le T} E[\sup_{t\le T} |D_{s_1}\ldots D_{s_n}\mathcal{E}_t|^{2q}] + \sup_{s_1,\ldots,s_n \le T} E[\sup_{t\le T} |D_{s_1}\ldots D_{s_n}\mathcal{E}_t^{-1}|^{2q}] \le C, \tag{3.27}$$

for some positive constant C independent of h which without loss of generality we assume is teh same as the one appearing in Lemma 3.2.

So we can take $Z_s^h = \mathcal{E}_s^{-1}$, $V_t^h = \mathcal{E}_t$, $p = 1$, γ and β as above, satisfying (3.16) and (3.17). From here, we conclude that (3.26) is bounded by some constant $A_1 > 0$ and

$$\left| E'\left(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \tilde{E}\left[\tilde{B}_s b(\tilde{Y}_{\eta(s)}, G(\tilde{Y}_{\eta(s)}; v_{\eta(s)})) \int_{\eta(s)}^s d\tilde{W}_r\right] dW_s\right)\right| \le A_1 th.$$

Repeating the same argument with all the other terms, we can find a proper constant A such that the thesis is satisfied. The proof for the case $|\tilde{E}(u(X_t, Y_t, \tilde{X}_t, \tilde{Y}_t)\tilde{\mathcal{E}}_t\tilde{U}_0(t))| \le Ath$ is similar, Indeed, as before, we may decompose $\tilde{\mathcal{E}}_t\tilde{U}_0(t)$ into its integral terms and apply the first part of Lemma 3.5 with a uniformly bounded smooth random function, as we observed in Remark 3.6 (b). □

We now prove the second step of the induction in the lemma that follows.

Lemma 3.9: *There exists a constant $R > 0$, independent of t, h, x and k such that*

$$|E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_k(t))| \le h \sum_{j=1}^{k+1} \frac{(Rt)^j}{j!}.$$

$$|\tilde{E}(u(\tilde{X}_t, \tilde{Y}_t)\tilde{\mathcal{E}}_t \tilde{U}_k(t))| \le h \sum_{j=1}^{k+1} \frac{(Rt)^j}{j!}.$$

Here $u : \Omega \times \mathbb{R}^2 \to \mathbb{R}$ is any smooth random measurable function with its first 4 derivatives bounded by M uniformly in Ω.

Proof: We proceed by induction. On the basis of the previous Lemma we are going to suppose that the result is true for k and prove the same inequalities for $k+1$. From the proof we will get the sufficient conditions to choose the constant R. Using (3.23), we see that the following inequality is satisfied

$$\begin{aligned}
&|E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_{k+1}(t))| \le |E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_0(t))| \\
&+ |E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \tilde{E}\left(a_y(Y_s, \eta^1_s) f_y(Y_s, \tilde{\xi}^1_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s)\right) ds| \\
&+ |E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} \tilde{E}\left(\beta_s b_y(Y_s, \eta^2_s) g_y(Y_s, \tilde{\xi}^2_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s)\right) ds| \\
&+ |E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b_y(Y_s, \eta^2_s) \tilde{E}\left(g_y(Y_s, \tilde{\xi}^2_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s)\right) dW_s)|.
\end{aligned}$$

By the previous lemma, the first term in the right hand side of the inequality is certainly less than or equal to Aht, hence let us focus our attention on the other two terms.

First of all let us rewrite the above inequality, by using the midpoint notation (3.12), therefore, recalling the definitions of β_s, ξ^1_s, ξ^3_s, ξ^5_s, η^1_s, and η^2_s, we have

$$\begin{aligned}
&|E'(\phi'_{\frac{h}{2}}(\xi^1_t + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_{k+1}(t))| \le Aht \\
&+ \int_0^1 \int_0^1 |E'(\phi'_{\frac{h}{2}}(Z^{\nu,\bar{W}}_t - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} a_y(Y_s, F^\epsilon_s)\tilde{E}(f_y(Y_s, \tilde{\xi}^1_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s))ds)| d\epsilon d\nu \\
&+ \int_0^1 \int_0^1 \int_0^1 |E'(\phi'_{\frac{h}{2}}(Z^{\nu,\bar{W}}_t - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b_x(Z^\lambda_s, G(X_s; u_s)) b_y(Y_s, G^\epsilon_s)\tilde{E}(g_y(Y_s, \tilde{\xi}^2_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s))ds)| d\lambda d\epsilon d\nu \\
&+ \int_0^1 \int_0^1 \int_0^1 |E'(\phi'_{\frac{h}{2}}(Z^{\nu,\bar{W}}_t - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b^2_y(Y_s, G^\epsilon_s) G'(Z^\rho_s; u_s)\tilde{E}(g_y(Y_s, \tilde{\xi}^2_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s))ds)| d\lambda d\epsilon d\nu \\
&+ \int_0^1 \int_0^1 |E'(\phi'_{\frac{h}{2}}(Z^{\nu,\bar{W}}_t - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b_y(Y_s, G^\epsilon_s)\tilde{E}(g_y(Y_s, \tilde{\xi}^2_s)\tilde{\mathcal{E}}_s \tilde{U}_k(s))dW_s)| d\epsilon d\nu
\end{aligned}$$

where we set $F^\epsilon_s = (1-\epsilon)F(X_s; u_s) + \epsilon F(Y_s; v_s)$, $G^\epsilon_s = (1-\epsilon)G(X_s; u_s) + \epsilon G(Y_s; v_s)$ and $Z^\lambda = X_t + \lambda(Y_t - X_t)$, for $0 \le \lambda, \nu, \epsilon \le 1$.

Let us notice that the functions

$$\beta_1(x_5,x_6)=a_y(x_6,(1-\epsilon)\int_{\mathbb{R}}f(x_5,z)d\mu_s(z)+\epsilon\int_{\mathbb{R}}f(x_6,z)dv_s(z))$$
$$\beta_2(x_5,x_6)=b_y(x_6,(1-\epsilon)\int_{\mathbb{R}}g(x_5,z)d\mu_s(z)+\epsilon\int_{\mathbb{R}}g(x_6,z)dv_s(z))$$
$$\beta_3(x_5,x_6)=b_x((1-\lambda)x_5+\lambda x_6,\int_{\mathbb{R}}g(x_5,z)d\mu_s(z))$$
$$\beta_4(x_5,x_6)=\int_{\mathbb{R}}g_x((1-\lambda)x_5+\lambda x_6,z)d\mu_s(z)$$
$$\beta_5(x_5,x_6)=\beta_2(x_5,x_6)\beta_3(x_5,x_6)+\beta_2^2(x_5,x_6)\beta_4(x_5,x_6)$$

all have derivatives up to order 4, uniformly bounded by a fixed constant depending on M, that we will denote with C_M.
We now want to apply Remark 3.6(a), taking $V_t^h=\mathcal{E}_t$, $Z_s^h=\mathcal{E}_s^{-1}\tilde{E}(f_y(Y_s,\tilde{\xi}_s^1)\tilde{\mathcal{E}}_s\tilde{U}_k(s))$ or $Z_s^h=\mathcal{E}_s^{-1}\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_k(s))$ and $\beta_s(x_1,\ldots,x_8)=\beta_i(x_1,x_2)$, $i=1,2,5$, so we have to verify that the hypotheses of Lemma 3.5 are satisfied. We have to find a bound for $\| Z_s^h\|_{n,q}$, for q large enough and $n\le 4$. For this, first note that $\mathcal{E}_t$ and $\mathcal{E}_s^{-1}$ verify (3.27), in the sense of Sobolev norms with respect only to W, just like in Remark 3.6(a).
Using the usual midpoint notation, our task is made equivalent to finding a bound for $||\mathcal{E}_s^{-1}\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))||_{n,q}$, where $Z^\tau=X_t+\tau(Y_t-X_t)$ for $0\le\tau\le 1$ and u is either f or g. It is important to observe that in this case $u_y(Y_s,\cdot)$ is a smooth random function independent of U_k and of $\tau\in[0,1]$.

By Hölder's inequality we have for $n\le 4$

$$\begin{aligned}\|\mathcal{E}_s^{-1}\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))\|_{n,q}&\le\|\mathcal{E}_s^{-1}\|_{n,q_1}\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))\|_{n,q_2}\\&\le C_1\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))\|_{n,q_2},\end{aligned}$$

with $\frac{1}{q_1}+\frac{1}{q_2}=\frac{1}{q}$. For example, consider the case when $n=2$. We derive our estimate only in this case, to keep the computations more understandable, for $n=3,4$ it is just a matter of considering also the Malliavin derivatives of order 3 and 4. By differentiating we obtain

$$\begin{aligned}D_r\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))&=D_rY_s\tilde{E}(u_{yx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))\\D_rD_u\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))&=D_rD_uY_s\tilde{E}(u_{yxx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))\end{aligned}$$

and consequently we have that

$$\begin{aligned}&\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))\|_{2,q_2}^{q_2}\le E(|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))|^{q_2})\\&+E\left[\left(\int_0^T|D_rY_s|^2|\tilde{E}(u_{yx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))|^2dr\right.\right.\\&\qquad\left.\left.+\int_0^T\int_0^T|D_rD_uY_s|^2|\tilde{E}(u_{yxx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))|^2dudr\right)^{\frac{q_2}{2}}\right].\end{aligned}$$

But $u_y(Y_s,\cdot)$ and $u_{yxx}(Y_s,\cdot)$ have derivatives uniformly bounded by M for all ω, therefore we see that the inductive hypotheses can be applied to conclude that

$$|\tilde{E}(u_{yxx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))| \le h\sum_{j=1}^{k+1}\frac{(Rt)^j}{j!}$$

$$|\tilde{E}(u_{yx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))| \le h\sum_{j=1}^{k+1}\frac{(Rt)^j}{j!}$$

$$|\tilde{E}(u_{y}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s))| \le h\sum_{j=1}^{k+1}\frac{(Rt)^j}{j!}$$

which implies $\| \tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_k(s)) \|_{2,q_2} \le h\sum_{j=1}^{k+1}\frac{(Rs)^j}{j!}(1+\| Y_s \|_{2,q_2}) \le Ch\sum_{j=1}^{k+1}\frac{(Rs)^j}{j!}$, because of Lemma 3.2. Summarizing, it is possible to find a constant $\bar{C}$ independent of all the parameters that depends polynomially on the constants M, T and the constant C in Lemma 3.2 such that
(3.28)

$$|E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_{k+1}(t))| \le Aht+\bar{C}C_1\int_0^t h\sum_{j=1}^{k+1}\frac{(Rs)^j}{j!}ds \le h(Rt+\sum_{j=1}^{k+1}\frac{(Rt)^{j+1}}{(j+1)!})$$

choosing $R \ge \max(A,\bar{C}C_1)$.

Similarly as in Lemma 3.8 one may prove that for a random function, independent of $\tilde{W}$ $u:\Omega\times\mathbb{R}^2\to\mathbb{R}$ with derivatives bounded by M uniformly in $\Omega\times\mathbb{R}^2$ one has that

$$|\tilde{E}(u(\tilde{Y}_t,\tilde{X}_t)\tilde{\mathcal{E}}_t\tilde{U}_{k+1}(t))| \le h(Rt+\sum_{j=1}^{k+1}\frac{(Rt)^{j+1}}{(j+1)!})$$

This concludes the proof. $\square$

We now want to establish the same result as Theorem 3.7, for the L^1 norm.

Theorem 3.10: *Under the same hypotheses of Theorem 3.1, the following inequality hold*

$$\int_{\mathbb{R}}|E(\phi_h(X_t-x)-\phi_h(Y_t-x))|dx \le Ch, \tag{3.29}$$

with C independent of h.

PROOF: Since the proof is a slight modification of that of Theorem 3.7, we are going to sketch it only.

By following exactly the same steps as before, we have by dominated convergence theorem that

$$\begin{aligned}\int_{\mathbb{R}} |E(\phi_h(X_t-x)-\phi_h(Y_t-x))|dx &= \int_{\mathbb{R}} |E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)(X_t-Y_t))|dx \\ &= \int_{\mathbb{R}} |E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_t)|dx \\ &= \lim_{k\to\infty}\int_{\mathbb{R}} |E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_k(t))|dx\,.\end{aligned}$$

On the other hand, studying the sequence of iterates we can see that

$$\int_{\mathbb{R}} |E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_0(t))|dx \tag{3.30}$$

is dominated by a sum of a finite number of terms of the type

$$\begin{aligned}\int_{\mathbb{R}}|\int_0^1\cdots\int_0^1\int_0^t\int_{\eta(s)}^s E'\Big(\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)[H_1(Z_t^{\nu,\bar{W}},M^{(1)}_{t,s,r}(\mu_1,\dots,\mu_l)) \\ +H_2(Z_t^{\nu,\bar{W}},M^{(2)}_{t,s,r}(\mu_1,\dots,\mu_l))]\Big)\,drdsd\mu_1\dots d\mu_l d\nu|dx,\end{aligned} \tag{3.31}$$

where $l\le 4$ is a fixed integer. In (3.30), we will study the term

$$E'''(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}\int_{\eta(s)}^s b(\tilde{Y}_{\eta(s)},G(\tilde{Y}_{\eta(s)};v_{\eta(s)}))d\tilde{W}_r\tilde{B}_sdW_s).$$

In this term we have $l=2$ and

$$\begin{aligned}M^{(1)}_{t,s,r} &= D_s\mathcal{E}_t\mathcal{E}_s^{-1}b(\tilde{Y}_{\eta(s)},\tilde{G}_s^{\mu_1})\mu_2g_{yy}(Y_s,\tilde{Y}_s^{\mu_2})\tilde{D}_r\tilde{Y}_sb(\tilde{Y}_{\eta(s)},G(\tilde{Y}_{\eta(s)};v_{\eta(s)}))\\ M^{(2)}_{t,s,r} &= D_sZ_t^{\nu,\bar{W}}\mathcal{E}_t\mathcal{E}_s^{-1}b(\tilde{Y}_{\eta(s)},\tilde{G}_s^{\mu_1})\mu_2g_{yy}(Y_s,\tilde{Y}_s^{\mu_2})\tilde{D}_r\tilde{Y}_sb(\tilde{Y}_{\eta(s)},G(\tilde{Y}_{\eta(s)};v_{\eta(s)}))\\ \tilde{G}_s^{\mu_1} &= (1-\mu_1)G(\tilde{Y}_{\eta(s)};v_{\eta(s)})+\mu_1G(\tilde{Y}_s;v_s)\\ \tilde{Y}_s^{\mu_2} &= (1-\mu_2)\tilde{Y}_{\eta(s)}+\mu_2\tilde{Y}_s.\end{aligned}$$

To simplify the notation we will omit some of the arguments for the rest of this proof. Since $\phi_{\frac{h}{2}}$ is a density function, it is positive, therefore (3.31) is bounded by

$$\begin{aligned}&\int_{\mathbb{R}}\int_0^1\int_0^1\int_0^1\int_0^t\int_{\eta(s)}^s E'''(\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)|H_1(Z_t^{\nu,\bar{W}},M_t^{(1)})+H_2(Z_t^{\nu,\bar{W}},M_t^{(2)})|drdsd\mu_1d\mu_2d\nu dx\\ &=\int_0^1\int_0^1\int_0^1\int_0^t\int_{\eta(s)}^s E'''\Big(\int_{\mathbb{R}}\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)dx\;|H_1(Z_t^{\nu,\bar{W}},M_t^{(1)})+H_2(Z_t^{\nu,\bar{W}},M_t^{(2)})|\Big)drdsd\mu_1d\mu_2d\nu\\ &=\int_0^1\int_0^1\int_0^1\int_0^t\int_{\eta(s)}^s E'''(|H_1(Z_t^{\nu,\bar{W}},M_t^{(1)})+H_2(Z_t^{\nu,\bar{W}},M_t^{(2)})|)drdsd\mu_1d\mu_2d\nu,\end{aligned}$$

where we used Fubini's theorem and the fact that $\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)$ integrates to one, as a density function, no matter what the value of $Z_t^{\nu,\bar{W}}$ is. Following the same steps as in the proof of Lemma 3.8, it is possible to prove that the integrand is bounded, independently of ν. Since no other quantity depends on x, we may conclude that

$$\int_{\mathbb{R}} |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_1 - x)\mathcal{E}_t U_0(t))|dx \leq Aht,$$

with A a constant independent of t, h and U_0. We then proceed by induction; using the definition of U_k (equation (3.23)), we arrive at the following inequality

$$\begin{aligned}
&\int_{\mathbb{R}} |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_k(t))|dx \\
\leq& Aht + \int_{\mathbb{R}} \int_0^1 \int_0^1 \int_0^1 \int_0^t |E'''[\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)\sum_{i=1}^4 Z_{k-1}^i(t,s)]ds| d\nu d\mu_1 d\mu_2 dx \\
\leq& Rht + \sum_{i=1}^4 \int_0^1 \int_0^1 \int_0^1 \int_0^t E'''[\int_{\mathbb{R}} \phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)dx |Z_{k-1}^i(t,s)|] ds d\nu d\mu_1 d\mu_2,
\end{aligned}$$

where, following the same notation as in the previous lemmas, we set

$$\begin{aligned}
Z_{k-1}^1(t,s) =& H\left(Z_t^{\nu,\bar{W}}, \mathcal{E}_t\mathcal{E}_s^{-1} a_y(Y_s, F_s^{\mu_2})\tilde{E}(f_y(Y_s, \tilde{Z}_s^{\mu_1})\tilde{\mathcal{E}}_s\tilde{U}_{k-1}(s))\right) \\
Z_{k-1}^2(t,s) =& H\left(Z_t^{\nu,\bar{W}}, \mathcal{E}_t\mathcal{E}_s^{-1}\beta_s b_y(Y_s, G_s^{\mu_2})\tilde{E}(g_y(Y_s, \tilde{Z}_s^{\mu_1})\tilde{\mathcal{E}}_s\tilde{U}_{k-1}(s))\right) \\
Z_{k-1}^3(t,s) =& H\left(Z_t^{\nu,\bar{W}}, D_s\mathcal{E}_t\mathcal{E}_s^{-1} b_y(Y_s, G_s^{\mu_2})\tilde{E}(g_y(Y_s, \tilde{Z}_s^{\mu_1})\tilde{\mathcal{E}}_s\tilde{U}_{k-1}(s))\right) \\
Z_{k-1}^4(t,s) =& H\left(Z_t^{\nu,\bar{W}}, D_s Z_t^{\nu,\bar{W}}\mathcal{E}_t\mathcal{E}_s^{-1} b_y(Y_s, G_s^{\mu_2})\tilde{E}(g_y(Y_s, \tilde{Z}_s^{\mu_1})\tilde{\mathcal{E}}_s\tilde{U}_{k-1}(s))\right).
\end{aligned}$$

As before, the density function integrates to 1 and $\sum_{i=1}^4 E'''(|Z_{k-1}^i(t,s)|) \leq Rh \sum_{j=1}^k \frac{(Rs)^j}{j!}$. By passing to the limit, our statement is proved. $\square$

We can pass to the next step in our procedure and consider the difference

$$E\left(\left|E(\phi_h(Y_t - x)) - \frac{1}{n}\sum_{j=1}^n \phi_h(Y_t^j - x)\right|\right),$$

where the Y^j are independent copies of Y. By using the Strong Law of Large Numbers we have that the difference converges to zero almost surely as $n \to \infty$ for fixed h. Moreover we can find the rate of convergence in $L^1(P)$, in fact

$$|E(\phi_h(Y_t - x)) - \frac{1}{n}\sum_{j=1}^n \phi_h(Y_t^j - x)| = |\frac{1}{n}\sum_{j=1}^n [E(\phi_h(Y_t - x)) - \phi_h(Y_t^j - x)]|$$

thus, by taking into account the independence of the copies, formula (3.7), Lemmas 3.2 and 3.3 and the boundedness of Φ_h, we obtain

$$
\begin{aligned}
&E\left[\Big(\frac{1}{n}\sum_{j=1}^{n}[E(\phi_h(Y_t-x))-\phi_h(Y_t^j-x)]\Big)^2\right] \\
\leq &\frac{1}{n^2}\sum_{j=1}^{n} E(\phi_h(Y_t-x))^2 = \frac{E'(\phi_{\frac{h}{4}}(Y_t+\sqrt{h}\bar{W}_{1/4}-x))}{2\sqrt{\pi h}n} \\
= &\frac{C}{\sqrt{h}n}|E'(\Phi_{\frac{h}{4}}(Y_t+\sqrt{h}\bar{W}_{1/4}-x)H(Y_t+\sqrt{h}\bar{W}_{1/4},1))| \\
\leq &\frac{C}{\sqrt{h}n}\|\gamma^{-1}_{Y_t+\sqrt{h}\bar{W}_{1/4}}\|_a\|Y_t+\sqrt{h}\bar{W}_{1/4}\|_{1,b}\leq\frac{C}{\sqrt{h}n}
\end{aligned}
$$

for some a, b positive constants and for all $x\in\mathbb{R}$. Consequently

$$
\sup_{x\in\mathbb{R}} E(|E(\phi_h(Y_t-x))-\frac{1}{n}\sum_{j=1}^{n}\phi_h(Y_t^j-x)|)\leq\frac{C}{\sqrt{\sqrt{h}n}}. \tag{3.32}
$$

Also we have

$$
\begin{aligned}
&\int_{\mathbb{R}} E\left(|E(\phi_h(Y_t-x))-\frac{1}{n}\sum_{j=1}^{n}\phi_h(Y_t^j-x)|\right)dx \\
\leq&\frac{1}{2\sqrt{\pi h}n}\int_{\mathbb{R}} E'(\phi_{\frac{h}{4}}(Y_t+\sqrt{h}\bar{W}_{\frac{1}{4}}-x))dx\leq\frac{C}{\sqrt{\sqrt{h}n}}.
\end{aligned} \tag{3.33}
$$

We are ready to proceed with our last step.

Theorem 3.11: *Under the same hypotheses of Theorem 3.1, for each $p>1$, there exist positive constants C_p and C, independent of x,t and h, such that*

$$
\sup_{x\in\mathbb{R}} E\left(\left|\frac{1}{n}\sum_{j=1}^{n}\phi_h(Y_t^j-x)-\frac{1}{n}\sum_{j=1}^{n}\phi_h(X_t^j-x)\right|\right)\leq C_p\frac{1}{h^{1-\frac{1}{2p}}\sqrt{n}} \tag{3.34}
$$

for $n=O(\frac{1}{h})^k$ for some $k>0$.

$$
\int_{\mathbb{R}} E\left(\left|\frac{1}{n}\sum_{j=1}^{n}\phi_h(Y_t^j-x)-\frac{1}{n}\sum_{j=1}^{n}\phi_h(X_t^j-x)\right|\right)dx\leq C\frac{1}{\sqrt{h}n}. \tag{3.35}
$$

PROOF: As usual, by applying the mean value theorem we may write

$$\phi_h(Y_t^j - x) - \phi_h(X_t^j - x) = \phi_h'(\rho_t^j - x)(Y_t^j - X_t^j),$$

with $\rho_t^j \in [Y_t^j; X_t^j]$. Following the same procedure as before, it is clear that the difference in (3.34) becomes

$$\begin{aligned}\frac{1}{n}E\left(\left|\sum_{j=1}^n[\phi_h(Y_t^j - x) - \phi_h(X_t^j - x)]\right|\right) &= \frac{1}{n}E\left(\left|\sum_{j=1}^n \phi_h'(\rho_t^j - x)(Y_t^j - X_t^j)\right|\right)\\ &\leq \frac{1}{n}\sum_{j=1}^n E\left(|\phi_h'(\rho_t^j - x)||Y_t^j - X_t^j|\right)\end{aligned}$$

In the case of (3.35) one can easily see that

$$\begin{aligned}\int_{\mathbb{R}} E\left(|\phi_h'(\rho_t^j - x)||Y_t^j - X_t^j|\right)dx &= E\left(\int_{\mathbb{R}}\frac{1}{h}|\rho_t^j - x|\phi_h(\rho_t^j - x)dx|Y_t^j - X_t^j|\right)\\ &= \sqrt{\frac{2}{\pi h}}E(|Y_t^j - X_t^j|).\end{aligned}$$

In the case of (3.31), with analogous notation as before, for $Z_t^{\nu,j} = (1-\nu)X_t^j + \nu Y_t^j$ we have

$$E\left(|\phi_h'(\rho_t^j - x)||Y_t^j - X_t^j|\right) = \int_0^1 E\left(|\phi_h'(Z_t^{\nu,j} - x)||Y_t^j - X_t^j|\right)d\nu. \tag{3.36}$$

Therefore, choosing $\frac{1}{p} + \frac{1}{q} = 1$, by Hölder's inequality, the integrand can be dominated as

$$E\left(|\phi_h'(Z_t^{\nu,j} - x)||Y_t^j - X_t^j|\right) \leq \parallel \phi_h'(Z_t^{\nu,j} - x)\parallel_p \parallel Y_t^j - X_t^j \parallel_q \tag{3.37}$$

Furthermore, by the properties of the Gaussian density

$$\begin{aligned}E(|\phi_h'(Z_t^{\nu,j} - x)|^p) &= \frac{1}{\sqrt{2\pi}^{p-1}}\frac{1}{\sqrt{p}h^{\frac{3p}{2}-\frac{1}{2}}}E(|Z_t^{\nu,j} - x|^p\phi_{\frac{h}{p}}(Z_t^{\nu,j} - x))\\ &\leq C\frac{1}{\sqrt{2\pi}^{p-1}}\frac{1}{\sqrt{p}h^p}E'(|\frac{V_t^{\nu,j} - x}{\sqrt{h}}|^p\phi_{\frac{1}{2p}}(\frac{V_t^{\nu,j} - x}{\sqrt{h}}))\\ &\leq C_p\frac{1}{h^p}\sqrt{h}\int_{\mathbb{R}}|y|^p\phi_{\frac{1}{2p}}(y)p_t^{\nu,j}(\sqrt{h}y + x)dy,\end{aligned}$$

where $p_t^{\nu,j}(y)$ denotes the density function of $V_t^{\nu,j} = Z_t^{\nu,j} + \sqrt{\frac{h}{2p}}\bar{W}_1$ The proof of (3.34) is finished once we prove that $\int_{\mathbb{R}}|y|^p\phi_{\frac{1}{2p}}(y)p_t^{\nu,j}(\sqrt{h}y + x)dy$ is bounded, for which it is enough to show the boundedness of $p_t^{\nu,j}(y)$. The link described at the beginning of section 2 between the density of a random variable and its Malliavin derivative (we are now considering the space $\Omega \times \bar{\Omega}$ with derivatives $D, \bar{D}$), can be applied here and we have that there exist positive constants a and b such that

$$p_t^{\nu,j}(y) = E'(1_{\{V_t^{\nu,j}>x\}} H(V_t^{\nu,j}, 1)) \leq \| \gamma_{V_t^{\nu,j}}^{-1} \|_a \| V_t^{\nu,j} \|_{1,b} < \infty. \tag{3.38}$$

By definition, it is clear that $\| V_t^{\nu,j} \|_{1,b} \leq \| Y_t^j \|_{1,b} + \| X_t^j - Y_t^j \|_{1,b} + \| \sqrt{\frac{h}{2p}} \bar{W}_1 \|_{1,b}$. Since, by Lemma 3.2, we know that $\| Y_t^j \|_{1,b}$ is finite, the whole question is reduced at evaluating $\| Y_t^j - X_t^j \|_b, \| X_t^j - Y_t^j \|_{1,b}$, for $b > 1$ and $\| \gamma_{V_t^{\nu,j}}^{-1} \|_a$. The first ones are proven in the next Lemma, while the second is shown in Lemma 3.13. Applying these results to (3.36), (3.37) and (3.38), we obtain our thesis. $\square$

Lemma 3.12: *For any $p > 1$, we have*

$$E(|Y_t^j - X_t^j|^p)^{\frac{1}{p}} \leq C\frac{1}{\sqrt{n}}, \qquad \| Y_t^j - X_t^j \|_{1,p} \leq C\frac{1}{\sqrt{n}}.$$

PROOF: We will only prove the first assertion for $p = 2$. The proofs of the second inequality and of the general case are similar. The difference $Y_t^j - X_t^j$ verifies the following equation

$$\begin{aligned} Y_t^j - X_t^j =& Y_{\eta(t)}^j - X_{\eta(t)}^j + [a(Y_{\eta(t)}^j, F(Y_{\eta(t)}^j; v_{\eta(t)})) - a(X_{\eta(t)}^j, F(X_{\eta(t)}^j; \bar{u}_{\eta(t)}))](t - \eta(t)) \\ &+ [b(Y_{\eta(t)}^j, G(Y_{\eta(t)}^j; v_{\eta(t)})) - b(X_{\eta(t)}^j, G(X_{\eta(t)}^j; \bar{u}_{\eta(t)}))](W_t^j - W_{\eta(t)}^j). \end{aligned}$$

We want to show that $Y_t^j - X_t^j$ is uniformly bounded in the L^2 norm. In order to show this, by virtue of the mean value theorem, we linearize the above equation. From now on, we denote by $Z^i = Y^i - X^i$, then

$$\begin{aligned} Z_t^i =& Z_{\eta(t)}^i + a_x(\xi_{\eta(t)}^1(i), F(Y_{\eta(t)}^i; v_{\eta(t)}))Z_{\eta(t)}^i(t - \eta(t)) \\ &+ a_y(X_{\eta(t)}^i, \theta_{\eta(t)}^1(i))[F(Y_{\eta(t)}^i; v_{\eta(t)}) - F(X_{\eta(t)}^i; \bar{u}_{\eta(t)})](t - \eta(t)) \\ &+ b_x(\xi_{\eta(t)}^2(i), G(Y_{\eta(t)}^i; v_{\eta(t)}))Z_{\eta(t)}^i(W_t^i - W_{\eta(t)}^i) \\ &+ b_y(X_{\eta(t)}^i, \theta_{\eta(t)}^2(i))[G(Y_{\eta(t)}^i; v_{\eta(t)}) - G(X_{\eta(t)}^i; \bar{u}_{\eta(t)})](W_t^i - W_{\eta(t)}^i), \end{aligned} \tag{3.39}$$

with $\theta_{\eta(t)}^1(i) \in [F(Y_{\eta(t)}^i; v_{\eta(t)}); F(X_{\eta(t)}^i; \bar{u}_{\eta(t)})]$, $\theta_{\eta(t)}^2(i) \in [G(Y_{\eta(t)}^i; v_{\eta(t)}); G(X_{\eta(t)}^i; \bar{u}_{\eta(t)})]$ and $\xi_{\eta(t)}^1(i), \xi_{\eta(t)}^2(i) \in [Y_{\eta(t)}^i; X_{\eta(t)}^i]$. By recalling the definition of F and G, and keeping in mind that the copies of X and those of Y are respectively identically distributed, we can write

$$F(Y^i_{\eta(t)}; v_{\eta(t)}) - F(X^i_{\eta(t)}; \bar{u}_{\eta(t)}) = \int_{\mathbb{R}} f(Y^i_{\eta(t)}, y) v_{\eta(t)}(dy) - \frac{1}{n}\sum_{j=1}^n f(X^i_{\eta(t)}, X^j_{\eta(t)})$$

$$= \frac{1}{n}\sum_{j=1}^n [E^j(f(Y^i_{\eta(t)}, Y^j_{\eta(t)})) - f(Y^i_{\eta(t)}, Y^j_{\eta(t)})] + \frac{1}{n}\sum_{j=1}^n [f(Y^i_{\eta(t)}, Y^j_{\eta(t)}) - f(X^i_{\eta(t)}, X^j_{\eta(t)})]$$

$$= \frac{1}{n}\sum_{j=1}^n [E^j(f(Y^i_{\eta(t)}, Y^j_{\eta(t)})) - f(Y^i_{\eta(t)}, Y^j_{\eta(t)})]$$

$$+ \frac{1}{n}\sum_{j=1}^n \{f_x(\lambda^{1,i}_{\eta(t)}, Y^j_{\eta(t)}) Z^i_{\eta(t)} + f_y(X^i_{\eta(t)}, \lambda^{2,j}_{\eta(t)}) Z^j_{\eta(t)}\},$$

with E^j denoting the expectation relative to W^j; similarly for the terms in G

$$G(Y^i_{\eta(t)}; v_{\eta(t)}) - G(X^i_{\eta(t)}; \bar{u}_{\eta(t)}) = \frac{1}{n}\sum_{j=1}^n [E^j(g(Y^i_{\eta(t)}, Y^j_{\eta(t)})) - g(Y^i_{\eta(t)}, Y^j_{\eta(t)})]$$

$$+ \frac{1}{n}\sum_{j=1}^n \{g_x(\lambda^{3,i}_{\eta(t)}, Y^j_{\eta(t)}) Z^i_{\eta(t)} + g_y(X^i_{\eta(t)}, \lambda^{4,j}_{\eta(t)}) Z^j_{\eta(t)}\},$$

where $\lambda^{1,i}_{\eta(t)}$, $\lambda^{3,i}_{\eta(t)} \in [Y^i_{\eta(t)}; X^i_{\eta(t)}]$ and $\lambda^{2,j}_{\eta(t)}$, $\lambda^{4,j}_{\eta(t)} \in [Y^j_{\eta(t)}; X^j_{\eta(t)}]$.

Hence Equation (3.39) becomes

$$\begin{aligned} Z^i_t = {} & \int_0^t [A^{i,i}_{\eta(s)} Z^i_{\eta(s)} + \sum_{j\neq i}^n A^{i,j}_{\eta(s)} Z^j_{\eta(s)}] ds + \int_0^t [B^{i,i}_{\eta(s)} Z^i_{\eta(s)} + \sum_{j\neq i}^n B^{i,j}_{\eta(s)} Z^j_{\eta(s)}] dW^i_s \\ & + \int_0^t C^i_{\eta(s)} ds + \int_0^t J^i_{\eta(s)} dW^i_s \end{aligned} \tag{3.40}$$

where for $i, j = 1, \ldots, n$

$$A^{i,i}_\cdot = a_x(\xi^1_\cdot(i), F(Y^i_\cdot; v_\cdot)) + a_y(X^i_\cdot, \theta^1_\cdot(i))(\frac{1}{n}\sum_{j=1}^n f_x(\lambda^{1,i}_\cdot, Y^j_\cdot) + \frac{1}{n} f_y(X^i_\cdot, \lambda^{2,i}_\cdot))$$

$$A^{i,j}_\cdot = a_y(X^i_\cdot, \theta^1_\cdot(i)) \frac{1}{n} f_y(X^i_{\eta(t)}, \lambda^{2,j}_{\eta(t)}) \qquad i \neq j$$

$$B^{i,i}_\cdot = b_x(\xi^2_\cdot(i), G(Y^i_\cdot; v_\cdot)) + b_y(X^i_\cdot, \theta^2_\cdot(i))(\frac{1}{n}\sum_{j=1} g_x(\lambda^{3,i}_{\eta(t)}, Y^j_{\eta(t)}) + g_y(X^i_\cdot, \lambda^{4,i}_\cdot))k$$

$$B^{i,j}_\cdot = -b_y(X^i_\cdot, \theta^2_\cdot(i)) \frac{1}{n} g_y(X^i_{\eta(t)}, \lambda^{4,j}_{\eta(t)}) \qquad i \neq j$$

$$C^i_\cdot = a_y(X^i_\cdot, \theta^1_\cdot(i)) \frac{1}{n}\sum_{j=1}^n [E^j(f(Y^i_\cdot, Y^j_\cdot)) - f(Y^i_\cdot, Y^j_\cdot)]$$

$$J^i_\cdot = b_y(X^i_\cdot, \theta^2_\cdot(i)) \frac{1}{n}\sum_{j=1}^n [E^j(g(Y^i_\cdot, Y^j_\cdot)) - g(Y^i_\cdot, Y^j_\cdot)],$$

form the entries of the matrices that we denote by A and B and of the vectors C and J. So equation (3.40) can be written in vector form as

$$Z_t^* = H_t^* + \int_0^t Z_{\eta(s)}^* dN_s^*, \tag{3.41}$$

where we are using $*$ to denote the transpose of a matrix and $dN_s^{i,j} = A_s^{i,j} ds + B_s^{i,j} dW_s^i$ and $dH_s^* = (C_{\eta(s)}^1 ds + J_{\eta(s)}^1 dW_s^1, \ldots, C_{\eta(s)}^n ds + J_{\eta(s)}^n dW_s^n)$. At the points of the partition, the process Z is given by $Z_{t_m}^* = \sum_{k=0}^{m-1} Z_{t_k}^* (N_{t_{k+1}}^* - N_{t_k}^*) + H_{t_m}^*$, which has unique solution (Protter (1990), page 271).

$$Z_{t_m}^* = U_{t_m}^* \sum_{k=0}^{m-1} (U^*)_{t_k}^{-1} \left[(H_{t_{k+1}}^* - H_{t_k}^*) - ([H^*, N^*]_{t_{k+1}} - [H^*, N^*]_{t_k}) \right], \tag{3.42}$$

where U^* and $(U^*)^{-1}$ are respectively the unique solutions of the matrix equations

$$U_t^* = I + \int_0^t U_{\eta(s)}^* dN_s^* \qquad (U^*)_t^{-1} = I - \int_0^t (d(N^* - [N^*, N^*])_s)(U^*)_{\eta(s)}^{-1}. \tag{3.43}$$

Let us remark that the entries of the matrices A and B are uniformly bounded, namely it is immediate to see that

$$|A^{i,i}|, \quad |B^{i,i}| \leq M^2 + M \quad \text{and} \quad |A^{i,j}|, \quad |B^{i,j}| \leq \frac{M^2}{n} \quad \text{for } i \neq j.$$

From (3.40), keeping in mind that $(\sum_{i=1}^n x_i)^2 \leq n \sum_{i=1}^n x_i^2$ and Jensen's inequality, we get

$$\begin{aligned} |Z_t^i|^2 \leq 6\{T \int_0^t & \left[|A_{\eta(s)}^{i,i}|^2 |Z_{\eta(s)}^i|^2 + (n-1) \sum_{j \neq i} |A_{\eta(s)}^{i,j}|^2 |Z_{\eta(s)}^j|^2 + |C_{\eta(s)}^i|^2 \right] ds \\ & + |\int_0^t B_{\eta(s)}^{i,i} Z_{\eta(s)}^i dW_s^i|^2 + |\int_0^t \sum_{j \neq i} B_{\eta(s)}^{i,j} Z_{\eta(s)}^j dW_s^i|^2 + |\int_0^t J_{\eta(s)}^i dW_s^i|^2\}. \end{aligned}$$

Taking the supremum over $[0,t]$ and the expectation, by employing Doob's inequality for martingales we finally obtain

$$\begin{aligned} E(\sup_{0 \leq s \leq t} |Z_s^i|^2) \leq 6TE(&\int_0^t \left[(M+M^2)^2 \sup_{0 \leq r \leq s} |Z_r^i|^2 + \frac{M^4}{n} \sum_{j \neq i} \sup_{0 \leq r \leq s} |Z_r^j|^2 + |C_{\eta(s)}^i|^2 \right] ds) \\ + 24E(&\int_0^t \left[(M+M^2)^2 \sup_{0 \leq r \leq s} |Z_r^i|^2 + \frac{M^4}{n} \sum_{j \neq i} \sup_{0 \leq r \leq s} |Z_r^j|^2 + |J_{\eta(s)}^i|^2 \right] ds) \end{aligned}$$

summarized into

$$\vartheta_i(t) \leq \int_0^t (K_1\vartheta_i(s) + \frac{K_2}{n}\sum_{j\neq i}\vartheta_j(s) + K_3 x_i(s))ds$$

where $\vartheta_i(t) = E(\sup_{0\leq s\leq t}|Z_s^i|^2)$, $x_1(s) = E(|J^i_{\eta(s)}|^2 + |C^i_{\eta(s)}|^2)$, $K_3 = 6T+24$, $K_2 = K_3M^4$ and $K_1 = K_3(M+M^2)^2$. Gronwall's inequality then implies

$$E(\sup_{0\leq s\leq t}|Z_s^i|^2) \leq \mathrm{e}^{K_1T}\int_0^t E[\frac{K_2}{n}\sum_{j\neq i}\sup_{0\leq r\leq s}|Z_r^j|^2 + K_3(|J^i_{\eta(s)}|^2 + |C^i_{\eta(s)}|^2)]ds,$$

(3.44)

$$\sum_{i=1}^n E(\sup_{0\leq s\leq t}|Z_t^i|^2) \leq \mathrm{e}^{K_4T}K_5\int_0^t\sum_{i=1}^n E(|J^i_{\eta(s)}|^2 + |C^i_{\eta(s)}|^2)ds$$

with $K_4 = \mathrm{e}^{K_1T}K_2$ and $K_5 = \mathrm{e}^{K_1T}K_3$. Consequently the problem is reduced to analyzing the vectors C and J. We can evaluate the last two expectations by the propagation of chaos. We focus our attention only on $E(\|C_{t_k}\|^2)$ ($\|\cdot\|$ here means the euclidean norm), as the other case is similarly carried out.

Since the sequence Y^i is formed by independent copies of the original process Y, also the processes $f(Y^i_\cdot, Y^j_\cdot)$ and $f(Y^i_\cdot, Y^l_\cdot)$ result conditionally independent, given Y^i, provided $j\neq l$, so for each i and each $r = t_k$, the following inequality is fulfilled

$$E(|C_r^i|^2) = E\{\,|a_y(X^i_{t_k}, \theta^1_r(i))\frac{1}{n}\sum_{j=1}^n[E^j(f(Y_r^i - Y_r^j)), f(Y_r^i, Y_r^j)]|^2\,\}$$

$$\leq\frac{M^2}{n^2}\{\,2E[\sum_{\substack{j,l\\ l<j}}[E^j(f(Y_r^i,Y_r^j)) - f(Y_r^i,Y_r^j)][E^l(f(Y_r^i,Y_r^l)) - f(Y_r^i,Y_r^l)]]$$

$$+\sum_{j=1}^n Var(f(Y_r^i,Y_r^j))\,\}$$

$$\leq\frac{M^2}{n^2}2E[E(\sum_{\substack{j,l\\ l<j}}[E^j(f(Y_r^i,Y_r^j)) - f(Y_r^i,Y_r^j)][E^l(f(Y_r^i,Y_r^l)) - f(Y_r^i,Y_r^l)]|Y_r^i)]$$

$$+\frac{4M^4}{n}$$

$$\leq\frac{M^2}{n^2}2E\{\sum_{\substack{j,l\\ l<j}}E[E^j(f(Y_r^i,Y_r^j))-f(Y_r^i,Y_r^j)|Y_r^i]E[E^l(f(Y_r^i,Y_r^l))-f(Y_r^i,Y_r^l)|Y_r^i]\}$$

$$+\frac{4M^4}{n} = 0 + \frac{4M^4}{n}.$$

Substituting in (3.44), we finally obtain

$$E(\sum_i \sup_{0\le t\le T}|Z_t^i|^2) \le \mathrm{e}^{K_4 T} K_5 \Gamma \sum_{k=0}^{n-1}(t_{k+1}-t_k)\frac{M^4}{n} = \frac{\Lambda}{n}t_m \le \frac{\Lambda}{n}T$$

for an appropriately chosen constants Γ, Λ and, of course, the same inequality holds for each component.

To prove the second statement, it remains to show that for all j,

$$\sum_{i=1}^n \int_0^T E(|D_s^i(Y_t^j - X_t^j)|^2)ds \le C\frac{1}{\sqrt{n}}.$$

Starting again from (3.41), it is possible to show that for each i the matrix process $(D_s^i Z_t^j) = X_{(i)}^j(s,t)$ verifies the linear matrix sde

$$X_{(i)}^*(s,t) = \tilde{K}_{(i)}(s,t) + \int_s^t X_{(i)}^*(s,r)d\tilde{N}_{(i)}^*(s,r),$$

where the matrices are given by

$$\begin{aligned} (d\tilde{N}_{(i)}^*(s,r))_{jk} &= D_s^i A_r^{j,k} dr \qquad \text{for } j \ne i \\ (d\tilde{N}_{(i)}^*(s,r))_{i,k} &= D_s^i A_r^{i,k} dr + D_s^i B_r^{i,k} dW_r^i \\ \tilde{K}_{(i)}^k(s,t) &= D_s^i H_t^k \qquad \text{for } k \ne i \\ \tilde{K}_{(i)}^i(s,t) &= D_s^i H_t^i + (Z_s^* B_s^*)^i. \end{aligned}$$

With computations similar to those shown before, it is possible to deduce an inequality analogous to (3.45) with the coefficients of $K_{(i)}$ in place of those of H, from which will descend the result by propagation of chaos and so we conclude the proo f. $\square$

We would like to remark that when we apply the inequality of Lemma 3.12 to our terms in Theorem 3.11 we have

$$\frac{1}{n}\frac{1}{\sqrt{2\pi h}}\sum_{j=1}^n E(|Y_t^j - X_t^j|^2)^{\frac{1}{2}} \le \frac{1}{\sqrt{2\pi h}}(\Lambda T)^{\frac{1}{2}}\frac{1}{\sqrt{n}}\frac{1}{n}n \le C\frac{1}{\sqrt{nh}},$$

giving the right order of convergence.

It remains to check the boundedness of the last factor

Lemma 3.13: *Let* $V_t^{\nu,j} = Z_t^{\nu,j} + \sqrt{\frac{h}{2p}}\tilde{W}_1$ *and* $n = O(\frac{1}{h})^k$ *for some* $k > 0$, *then the following holds*

$$\sup_{h\in(0,1]}\sup_{\nu\in[0,1]} \|\gamma_{V_t^{\nu,j}}^{-1}\|_p < \infty \textit{ for all } p \in \mathbb{N} \textit{ and } t \in (0,T].$$

PROOF: Let X^j denote the unique strong solution to (1.1) when the stochastic equation is driven by W^j. The three main points that one needs to check in order to prove the boundedness of the Malliavin covariance matrix are

(i) $\sup_{\nu\in[0,1]} \| V_t^{\nu,j} - X_t^j \|_{1,p} \leq C(\frac{1}{\sqrt{n}} + \sqrt{h})$;

(ii) $\| \gamma_{X_t^j}^{-1} \|_p < \infty$ for all $p \in \mathbb{N}$;

(iii) $\| \gamma_{V_t^{\nu,j}}^{-1} \|_p \leq Ch^{-1}$.

The first and the third inequality come directly from Lemma 3.3, in particular using formula (3.7), while the second one was proven in Lemma 2.2 for the process X, but clearly the same is true for the copies. □

4. Proof of Theorem 3.1

This brief section is dedicated to gather all the results that we exposed in the previous ones and to finally obtain the proof of Theorem 3.1.

The statement of Theorem 3.1 deals with both density and distribution functions, but as we announced, we focus our attention only on the proof for the first ones, for which we have laid out all the necessary results. To get the same conclusion in the case of distributions, the whole procedure should be reconstructed, but we will just describe it briefly.

For the densities, we first consider the L^1 norm (3.4)

$$
\begin{aligned}
\int_{\mathbb{R}} E\left(|p_t(x) - \frac{1}{n}\sum_{j=1}^n \phi_h(X_t^j - x)| \right) dx &\leq \int_{\mathbb{R}} |p_t(x) - E(\phi_h(X_t - x))| dx \\
&+ \int_{\mathbb{R}} |E(\phi_h(X_t - x) - \phi_h(Y_t - x))| dx \\
&+ \int_{\mathbb{R}} E[|E(\phi_h(Y_t - x)) - \frac{1}{n}\sum_{j=1}^n \phi_h(Y_t^j - x)|] dx \\
&+ \int_{\mathbb{R}} E[|\frac{1}{n}\sum_{j=1}^n \phi_h(Y_t^j - x) - \frac{1}{n}\sum_{j=1}^n \phi_h(X_t^j - x)|] dx \\
&\leq C(h + \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{nh}}).
\end{aligned}
$$

The above bounds follow from Lemma 3.4, Theorem 3.10, (3.33) and Theorem 3.11. The analogous result (3.6), when adopting the norm of the supremum follows by applying instead Lemma 3.4, Theorem 3.7, (3.32) and Theorem 3.11.

Consider now the proofs for distribution functions (3.3):

$$
\begin{aligned}
\int_{\mathbb{R}} E\left[|u(t,x) - \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_t^j \le x\}}|\right] dx &\le \int_{\mathbb{R}} |E(1_{\{X_t \le x\}} - 1_{\{Y_t \le x\}})| dx \\
&+ \int_{\mathbb{R}} |E(1_{\{Y_t \le x\}} - \frac{1}{n}\sum_{j=1}^{n} 1_{\{Y_t^j \le x\}})|\, |dx \\
&+ \int_{\mathbb{R}} E\left[|\frac{1}{n}\sum_{j=1}^{n} 1_{\{Y_t^j \le x\}} - \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_t^j \le x\}}|\right] dx \\
&= A_1 + A_2 + A_3
\end{aligned}
$$

Let us consider the quantity A_1, for this one has to prove that there exits a positive constant C independent of $\epsilon \in (0,1]$ and h such that

$$
\begin{aligned}
|E(1_{\{X_t \le x\}} - \Phi_\alpha(x - X_t))| &\le C\epsilon \\
|E(\Phi_\epsilon(x - X_t) - \Phi_\epsilon(x - Y_t))| &\le Ch \\
|E(\Phi_\epsilon(x - Y_t) - 1_{\{Y_t \le x\}})| &\le C\epsilon.
\end{aligned}
$$

The first and third assertion are proven by the same argument as in Lemma 3.4 (note that we know that $\gamma^{-1}_{Y_t + \sqrt{\epsilon} \bar{W}_{\frac{1}{2}}} \in \cap_{p>1} L^p(\Omega)$ by taking $\nu = 0$ in Lemma 3.3), while the second one is proven along the lines of the proof of Theorem 3.7.

The quantity A_2 can be analyzed in the same way as we showed (3.33), while for A_3, we have

$$
\begin{aligned}
\int_{\mathbb{R}} E\left(|\frac{1}{n}\sum_{j=1}^{n} 1_{\{Y_t^j \le x\}} - \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_t^j \le x\}}|\right) dx &\le \frac{1}{n}\sum_{j=1}^{n} E(\int_{\mathbb{R}} |1_{\{Y_t^j \le x\}} - 1_{\{X_t^j \le x\}}| dx) \\
&\le \frac{1}{n}\sum_{j=1}^{n} E(|Y_t^j - X_t^j|) \le \frac{C}{\sqrt{n}}.
\end{aligned}
$$

This finishes the proof of (3.3). For (3.5) the proof is similar to the one for (3.4). $\square$

Conclusions

In this work, we have analyzed the rate of convergence of a particle method introduced by Bossy and Talay in order to approximate the solution to the Mc Kean-Vlasov equation and we showed that the rate of convergence is faster than the rate obtained by the authors in their article. On the other hand, the rate of convergence obtained here seems to match their simulations run in the particular case of the Burgers equation.

We also analyzed the rate of convergence when approximating the marginal densities of the solution. In order to carry out the necessary calculations we had to study the existence and smoothness of these densities.

The problem of obtaining the optimal rate of convergence for the Burgers equations is still open and the authors hope the method developed here might apply, if properly adapted, also to this case.

Some straightforward generalizations of the above results were not included in our exposition for reasons of space. For instance, it is not difficult to consider the case when also the initial random variable has to be approximated or when the measurements of the error is done through the variances (i.e. $L^2(\Omega)$) rather than through the expectations. Yet another generalization is to consider approximations of the type ϕ_ϵ rather than ϕ_h; if $\epsilon = O(h^r)$ for some $r > 0$ a similar analysis can be carried out.

Finally we remark that the condition $n = O(\frac{1}{h})^k$ for some $k > 0$ in Theorem 3.1 (used to obtain Lemma 3.13) is merely technical rather than restrictive, since k can be chosen freely.

Acknowledgements: This work was completed while the first author's visit at Purdue University and he thanks the Mathematics Department for the hospitality. During this period the research of the first author was partially supported by the CNR-NATO grant n. 215.29. The research of the second author was completed while visiting the Department of Mathematics of the University of Tokyo. He thanks them for their hospitality and his research was also supported by a DGICYT grant.

References

[BT1] M. Bossy and D. Talay, *A stochastic particle method for the McKean-Vlasov and the Burgers equation*, Math. Comp. **66** **217** (1997), 157–192.

[BT2] M. Bossy and D. Talay, *Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation*, Annals of Applied Prob. **6** **3** (1996), 818-861.

[C] A.L. Chorin, *Vortex methods and vortex statistics.*, In:Lecture for Les Houches summer school of theoretical physics., Lawrence Berkeley Laboratory Prepublications (1993).

[G] J. Gärtner, *On the Mckean-Vlasov limit for interacting diffusions*, Math. Nachr. **137** (1988), 197–248.

[H1] O.H. Hald, *Convergence of random methods for a reaction diffusion equation*, SIAM J. Sci. Stat. Comput. **2** (1981), 85–94.

[H2] O.H. Hald, *Convergence of a random method with creation of vorticity.*, SIAM J. Sci. Stat. Comput. **7** (1986), 1373–1386.

[KO] A. Kohatsu-Higa and S. Ogawa, *Weak rate of convergence for an Euler scheme of Nonlinear SDE's*, Monte Carlo Methods and Its Applications **3** (1997), 327–345.

[KP] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer Verlag, Berlin, Heidelberg, New York, 1992.

[M] S. Meleard, *Asymptotic Behaviour of some particle systems; McKean-Vlasov and Boltzmann models*, In: Probabilistic Models for Nonlinear Partial Differential Equations, Lecture Notes in Mathematics **1627** (1996), Springer Verlag, 42–95.

[N1] D. Nualart, *Malliavin Calculus and Related Topics*, Probability and its Applications, Springer Verlag, 1996.

[N2] D. Nualart, *Analysis on Wiener space and anticipating stochastic calculus*, In: Ecole d'été de Probabilités de Saint-Flour XXV (1995), Lecture Notes in Mathematics **1690** (1998 pages 123–227), Springer Verlag.

[P] P. Protter, *Stochastic Integration and Differential Equations - a New Approach*, Springer Verlag, 1990.

[Ta] D. Talay, *Probabilistic numerical Methods for Partial Differential Equations: Elements of Analysis*, in Probabilistic Models for Nonlinear Partial Differential Equations, Lecture Notes in Mathematics **1627** (1996), Springer Verlag, 148–196.

[T] S. Taniguchi, *Applications of Malliavin Calculus to time-dependent Systems of Heat Equations*, Osaka J. Math. **22** (1985), 307–320.

[S] D. W. Stroock, *Some Applications of Stochastic Calculus to Partial Differential Equations*, Lecture Notes in Mathematics **976** (1983), Springer Verlag, 268–382.

E-mail address: antonf@mat.uniroma1.it kohatsu@upf.es