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

Fabio Antonelli

Università di Roma "La Sapienza"
Dipartimento di Matematica,
Piazzale A. Moro, 2
00185 Rome - ITALY.

Arturo Kohatsu-Higa

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

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.

ABSTRACT. This paper studies the rate of convergence of an appropriate discretization scheme of the solution of the Mc Kean - Vlasov equation introduce d 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 st ep 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 and in the sup norm is of the order of $\frac{1}{\sqrt{n}} + h$, while for the densities is of the o rder $h + \frac{1}{\sqrt{nh}}$. This result is obtained by carefully employing techniques of Malliavin Calculus.
Keywords: Mc Kean - Vlasov equation, Malliavin Calculus.
Journal of Economic Literature Classification System:C15

1991 Mathematics Subject Classification. Primary: 60H10, 60K35 Secondary: 34F05 90A12

1. Introduction

In a series of recent articles (see [BT1] , [BT2], [T]), 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 towards the solution of the McKean - Vlasov limit with a rate at 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

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

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]).

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

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

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 $t \in [0,T]$

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

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

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

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 the hypotheses on the coefficients as well as those on the initial density function. We establish this result in Section 2 and it is quite related to similar ones obtained by Florchinger (see [F]), who was interested in an application of Malliavin Calculus to filtering theory, which required the study of the smoothness of densities for time dependent systems.

The main difference between Florchinger's results and ours is that we do not require any boundedness for the coefficients, since we show that a global Hölder property in t 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 one were to introduce an uniform Hörmander type condition on the coefficients, 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 an uniform Hörmander type 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

$$E\|E(\delta_x(X_t)) - \frac{1}{n}\sum_{j=1}^{n} \phi_\epsilon(X_t^j - x)\|_{L^1()}$$

$$\leq \|E(\delta_x(X_t)) - E(\delta_x(X_t^1))\|_{L^1()} \; + E\|E(\delta_x(X_t^1)) - \frac{1}{n}\sum_{j=1}^{n} \phi_\epsilon(X_t^j - x)\|_{L^1()}$$

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. But 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.

Our results can be easily written in the multidimensional case, but to keep notations and proofs simple, we decided to restrict ourselves to only one dimension.

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). This is where we modify Florchinger's results to our needs. In Section 3 we establish our 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 one dimensional Brownian motion, W, is defined. We consider the equation

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

where ξ is an $\mathcal{F}_0$-measurable random variable, 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 $\int \zeta(x,y)d\mu_s(y)$ ($\zeta = f$, g, respectively), where μ_s indicates the distribution of X_s. Lastly, the functions

$$b,\, a\ : \mathbb{R}^2 \longrightarrow \mathbb{R} \qquad f,\, g\ : \mathbb{R}^2 \longrightarrow \mathbb{R}$$

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.

Assumptions:

(H1) There exist an integer m and a positive constant c, such that

$$\sum_{i=0}^{m}\sum_{v\in I_i} v(0,\xi)^2 \geq c > 0 \quad a.e.,$$

where the sets I_i are given by

$$I_0 = \{\bar{b}\},\ \dots\ ,\ I_n = \{\ [\bar{b},v],\ [\bar{a},v],\ v\in I_{n-1}\}$$

and ($[\cdot,\cdot]$ denotes Lie bracket. In this context, the coefficients are to be understood as vector fields, that is $\bar{b}(t,x) = \bar{b}(t,x)\frac{d}{dx}$).

(H2) The function b is 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

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

With this notation, the hypothesis corresponding to (H-1) in Theorem 1.1 should be $\bar{b}(0,x) \geq c > 0$, for all $x \in \mathbb{R}$ and it is clear that (H1) requires much less than this. 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).

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

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

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 $n\in\mathbb{N}$ and $f\in C_b^{\infty}(\mathbb{R}^{nm})$, we can complete this space under the norm $\|\cdot\|_{1,p}$ given by

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

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 $j=1,\dots,m$, 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

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

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

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

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

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

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

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

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

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

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 $\gamma_F^{-1}DF$ 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).

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* $p > 1$

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

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. We remind the reader that, from now on, we will assume all our quantities to be one dimensional and we will use the multidimensional notation for Malliavin Calculus only later on, when needed. In this setting the Malliavin covariance matrix clearly reduces to $\gamma_F = ||DF||^2_{L^2(T)}$.

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

PROOF: As a starting point, let us remark that equation (2.1), with the new notation, can be rewritten as

$$X_t = \xi + \int_0^t \bar{a}(s, X_s)ds + \int_0^t \bar{b}(s, X_s) \cdot dW_s, \tag{2.2}$$

where in fact the coefficients are time dependent. Moreover, because of (H0), $\bar{a}$ and $\bar{b}$ are smooth in space with bounded derivatives (hence they are also globally Lipschitz) and they are globally Hölder of order 1/2 in time. Indeed, for $s, t \in [0, T]$,

$$\begin{aligned} |\bar{b}(t,x) - \bar{b}(s,x)| &= |b(x, \int g(x,y)d\mu_t(y)) - b(x, \int g(x,y)d\mu_s(y))| \\ &\leq M| \int g(x,y)d\mu_t(y) - \int g(x,y)d\mu_s(y)| = M|E(g(x,X_t) - g(x,X_s))| \\ &\leq M^2 E(|X_t - X_s|) \leq M^2 E(|X_t - X_s|^2)^{\frac{1}{2}} \leq C|t-s|^{\frac{1}{2}}, \end{aligned}$$

the same applies to $\bar{a}$. Renaming M properly, we may assume without loss of generality that there exists a common constant M bounding both the derivatives and the Hölder's constant C.

If it is (H2) to hold, one can follow exactly the same proof as in [F] (Theorem 1.2.7), with only two slight modifications and for this reason we refer the reader to Florchinger's paper, indicating only what formal changes are needed.

The first difference lies in the fact that Florchinger considers both coefficients to be bounded, while here we are taking only the diffusion one as such. By examining carefully his proof, it can be realized that the boundedness of the coefficients is required in order to define a certain constant, denoted K_X in inequality (1.17) of page 208

$$K_X = \sup_{r\in\{0,\ldots,m\}} \sup_{(t,x)\in[0,T]\times} (|X_r(t,x)| + |DX_r(t,x)|,$$

where X_0 is the drift coefficient, while $X_1, \ldots, X_m$ are the diffusion ones. In truth, the hypothesis on the drift coefficient is redundant, in fact the proof involves only the diffusion ones, therefore it is holds if we simply drop the requirement for $r = 0$. In our case this amounts to using the constant

$$K_b = \sup_{(t,x)\in[0,T]\times} (|\bar{b}(t,x)| + |D\bar{b}(t,x)|),$$

which is certainly bounded, by virtue of (H2).

When b is bounded the second difference becomes minor. This difference consists of changing the initial condition in [F] from a deterministic to a random one, the same argument goes through, thanks to the integrability of ξ. One then proceeds along the same lines and proves the smoothness of the densities of the solution.

Let us now assume (H3). Here the fact that ξ is random creates a significant problem as b is not bounded anymore. In fact bounds for ξ will show up in almost all the expressions (through the presence of τ) when using the Lipschitz property of b. This argument is somewhat involved, although the basic idea may be simple. As quite a few changes are required, we explain the technique more at length.

In this case $\bar{a}$ and $\bar{b}$ may be not bounded, but we are able to compensate this drawback by the fact that they are Hölder uniformly on the whole space, while in [F] this property is satisfied only locally.

Let ζ_t denote the derivative of the stochastic flow associated with equation (2.2) and ζ^{-1} the inverse flow, then both $\sup_{s\leq t}|\zeta_t(\xi)|$ and $\sup_{s\leq t}|\zeta_t^{-1}(\xi)| \in \bigcap_{p\geq 1} L^p$ and we can write the Malliavin derivative as $D_s X_t = \zeta_t(\xi)\zeta_s^{-1}(\xi)\bar{b}(s,X_s)$ (see [N1], page 109).

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

$$\gamma_{X_t}^{-1} = [(\zeta_t(\xi))^2 \int_0^t (\zeta_s^{-1}(\xi))^2 \bar{b}(s,X_s)^2 ds]^{-1} \in \bigcap_{p\geq 1} L^p.$$

Using the L^p boundedness of $\zeta_t(\xi)$ and Lemma 2.3.1 in [N1], this amounts to show that for all $p \geq 2$ there exists $\epsilon_0(p)$ such that for all $\epsilon \leq \epsilon_0(p)$

$$P\left(\int_0^t (\zeta_s^{-1}(\xi))^2 \bar{b}(s,X_s)^2 ds < \epsilon\right) \leq \epsilon^p. \tag{2.3}$$

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

$$\begin{aligned} P(\int_0^t (\zeta_s^{-1}(\xi))^2 \bar{b}(s,X_s)^2 ds < \epsilon) &\leq P(\int_0^t (\zeta_s^{-1}(\xi))^2 \bar{b}(s,X_s)^2 ds < \epsilon,\, |\xi| \leq \tau) + P(|\xi| \geq \tau) \\ &= p_1 + p_2. \end{aligned}$$

Clearly for p_2 we will use hypothesis (H3), which gives $P(|\xi| > \tau) \leq C\exp(-\alpha\tau^\beta)$ for $\tau \geq \rho$. 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 $|\xi| \leq \tau$, one also has that

$$\sum_{i=0}^m \sum_{v\in I_i} v(s,y)^2 \geq \frac{3c}{4} > 0, \qquad \text{a.e.} \tag{2.4}$$

for $|y-\xi| \leq R$ and $s \leq R^2$, where we can chose

$$R = \frac{c}{4(L(1+\tau) + \sqrt{L^2(1+\tau)^2 + cL})}.$$

(Here $L = 2^{m+1} \max_{v\in I_i, i=0,\ldots,m} K_v(K_v + |v(0,0)|)$, where K_v denotes the maximum of the Lipschitz and the Hölder constants of v for $v \in I_i$, $i = 0, ..., m$).

Moreover the following property holds:

$$R \in [\frac{c}{4L(2(1+\tau)+c)}, \frac{c}{4L}], \tag{2.5}$$

for some universal constant c. Let us fix the quantities $r = R/2$ and $t^0 = R^2/2$; clearly, without loss of generality, we can assume that $t^0 \le t$, moreover we define the stopping time $\sigma = \inf\{s : |X_s - \xi| \ge r \text{ or } |\zeta_s^{-1}(\xi) - 1| \ge \frac{1}{2}\} \wedge t^0$. To subdivide further p_1, we take a partition $0 = t_0 < t_1 < ... < t_N = t^0$ with mesh $|t_{i+1} - t_i| < \delta := \epsilon^4$ and $N = [\frac{t^0}{\delta}] + 1$, so that we can write

$$\begin{aligned}
p_1 &\le P(\int_0^\sigma \sum_{i=0}^N (\zeta_s^{-1}(\xi)\bar{b}(s,X_s))^2 1_{[t_i,t_{i+1})}(s)ds < \epsilon, |\xi| \le \tau) \\
&\le P\left(\int_0^\sigma \sum_{i=0}^N |(\zeta_s^{-1}(\xi)\bar{b}(s,X_s))^2 - (\zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s}))^2| 1_{[t_i,t_{i+1})}(s)ds > \frac{\epsilon}{2}, |\xi| \le \tau\right) \\
&+ P\left(\int_0^\sigma \sum_{i=0}^N (\zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s}))^2 1_{[t_i,t_{i+1})}(s)ds < \frac{3\epsilon}{2}, |\xi| \le \tau\right) \\
&= p_{11} + p_{12},
\end{aligned}$$

where $X_{t_i,s}$ stands for the process defined as the solution of

$$X_{t_i,s} = X_{t_i} + \int_{t_i}^s \bar{a}(t_i, X_{t_i,u})du + \int_{t_i}^s \bar{b}(t_i, X_{t_i,u}) \cdot dW_u$$

and similarly $\zeta_{t_i,s}$ stands for the derivative of the flow associated to the above equation.

The rest of the proof consists of proving the following two assertions

$$p_{11} \le O(\epsilon^q) \tag{A.1}$$

$$p_{12} \le \tau^{q/\nu} O(\epsilon^q), \quad \text{for some fixed } \nu > 0 \text{ and any } \quad q > 0,\ \tau \ge \rho. \tag{A.2}$$

To prove (A.1), one has to estimate for $t_i \le s \wedge \sigma$ the difference under the integral sign

$$\begin{aligned}
&|\,(\zeta_s^{-1}(\xi)\bar{b}(s,X_s))^2 - (\zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s}))^2\,| \\
=& |\,\zeta_s^{-1}(\xi)\bar{b}(s,X_s) - \zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s})\,|\,|\,\zeta_s^{-1}(\xi)\bar{b}(s,X_s) + \zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s})\,| \\
\le& \left\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|\,|\bar{b}(s,X_s)| + |\zeta_{t_i,s}^{-1}(\xi)|\,|\bar{b}(s,X_s) - \bar{b}(t_i,X_{t_i,s})|\right\} \\
&\cdot \left\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|\,|\bar{b}(s,X_s)| + |\zeta_{t_i,s}^{-1}(\xi)|\,|\bar{b}(s,X_s) + \bar{b}(t_i,X_{t_i,s})|\right\} \\
\le& \left\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|\ \left[|\bar{b}(s,X_s) - \bar{b}(s,\xi)| + |\bar{b}(s,\xi) - \bar{b}(s,0)| + |\bar{b}(s,0)|\right]\right. \\
&\left. + |\zeta_{t_i,s}^{-1}(\xi)|\left[|\bar{b}(s,X_s) - \bar{b}(s,X_{t_i,s})| + |\bar{b}(s,X_{t_i,s}) - \bar{b}(t_i,X_{t_i,s})|\right]\right\} \\
&\cdot \left\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|\ \left[|\bar{b}(s,X_s) - \bar{b}(s,\xi)| + |\bar{b}(s,\xi) - \bar{b}(s,0)| + |\bar{b}(s,0)|\right]\right. \\
&\left. + |\zeta_{t_i,s}^{-1}(\xi)|\left[|\bar{b}(s,X_s) - \bar{b}(s,X_{t_i,s})| + |\bar{b}(s,X_{t_i,s}) - \bar{b}(t_i,X_{t_i,s})| + 2|\bar{b}(t_i,X_{t_i,s})|\right]\right\}.
\end{aligned}$$

But the function $\bar{b}$ is continuous, hence $\sup_{s\in[0,T]} |\bar{b}(s,0)| = c_0 < \infty$; besides it is globally Lipschitz in x and uniformly Hölder in t of order $1/2$ with constant M in the whole space. Taking into account all these factors and the running hypotheses (mainly, $s \leq \sigma$), we have

$$\begin{aligned} |\bar{b}(s,\xi) - \bar{b}(s,0)| &\leq M|\xi| \\ |\bar{b}(s,X_s) - \bar{b}(s,\xi)| &\leq M|X_s - \xi| \leq Mr \\ |\bar{b}(s,X_s) - \bar{b}(s,X_{t_i,s})| &\leq M|X_s - X_{t_i,s}| \\ |\bar{b}(s,X_{t_i,s}) - \bar{b}(t_i,X_{t_i,s})| &\leq M|s-t_i|^{\frac{1}{2}} \leq M\delta^{\frac{1}{2}} = M\epsilon^2. \end{aligned}$$

By choosing $k_1 = \max\{c_0, M\}$, we may conclude

$$\begin{aligned} &|(\zeta_s^{-1}(\xi)\bar{b}(s,X_s))^2 - (\zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s}))^2| \\ \leq& k_1^2\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|(r+|\xi|+1) + |\zeta_{t_i,s}^{-1}(\xi)|(|X_s - X_{t_i,s}| + \epsilon^2)\} \\ &\cdot\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|(r+|\xi|+1) + |\zeta_{t_i,s}^{-1}(\xi)|(3|X_s - X_{t_i,s}| + \epsilon^2 + 2(r+|\xi|+1))\} \\ \leq& c_1\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|(r+|\xi|+1) + |\zeta_{t_i,s}^{-1}(\xi)|(|X_s - X_{t_i,s}| + \epsilon^2)\} \\ &\cdot\{|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|(r+|\xi|+1) + |\zeta_{t_i,s}^{-1}(\xi)|(|X_s - X_{t_i,s}| + \epsilon^2 + r + |\xi| + 1)\}, \end{aligned}$$

where c_1 was chosen as $3k_1^2$. This estimate is the first difference between the argument presented in [F] and ours. Applying it, one obtains (from now on, we omit the dependence in ξ and we denote by J_i the intervals $[t_i \wedge \sigma, t_{i+1} \wedge \sigma)$)

$$\begin{aligned} p_{11} \leq \sum_{i=0}^{N} P\Bigg(c_1 \int_{J_i} &\{|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|(r+|\xi|+1) + |\zeta_{t_i,s}^{-1}|(|X_s - X_{t_i,s}| + \epsilon^2)\}\cdot \\ &\{|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|(r+|\xi|+1) + |\zeta_{t_i,s}^{-1}|(|X_s - X_{t_i,s}| + \epsilon^2 + r + |\xi| + 1)\}ds > \frac{\epsilon}{2N}, |\xi| \leq \tau\Bigg). \end{aligned}$$

To evaluate the right side of the previous inequality, for a fixed $K \in \mathbb{R}^+$, we introduce the following sets

$$\begin{aligned} A_i &= \{\sup_{s\in J_i} |\zeta_s^{-1} - \zeta_{t_i,s}^{-1}| \leq K\epsilon\}, \qquad B_i = \{\sup_{s\in J_i} |X_s - X_{t_i,s}| \leq K\epsilon\}, \\ A_i^\tau &= A_i \cap \{|\xi| \leq \tau\}, \qquad B_i^\tau = B_i \cap \{|\xi| \leq \tau\}. \end{aligned}$$

Just for briefness, in the next few lines we will indicate with L_s the process under the integral sign in the estimate of p_{11}. With this notation we have

$$\begin{aligned} p_{11} \leq& \sum_{i=0}^{N} P\left(c_1 \int_{J_i} L_s ds > \frac{\epsilon}{2N}, |\xi| \leq \tau\right) \leq \sum_{i=0}^{N}\left[P\left(\{c_1 \int_{J_i} L_s ds > \frac{\epsilon}{2N}\} \cap (A_i^\tau \cap B_i^\tau)\right)\right] \\ &+ \sum_{i=0}^{N}\left[P\left(\{c_1 \int_{J_i} L_s ds > \frac{\epsilon}{2N}\} \cap (A_i^c \cup B_i^c) \cap \{|\xi| \leq \tau\}\right)\right]. \end{aligned}$$

Let us consider the first probability; by construction, on $(A_i^\tau \cap B_i^\tau)$ we have

$$|\xi| \le \tau \,, \qquad \sup_{s\in J_i} |\zeta_s^{-1} - \zeta_{t_i,s}^{-1}| \le K\epsilon \quad \text{and} \quad |X_s - X_{t_i,s}| \le K\epsilon.$$

Because of the way we chose the partition, we know that $|J_i| \le \delta = \epsilon^4$; besides, all its points are less than or equal to σ, so necessarily $|\zeta_s^{-1}| \le \frac{3}{2}$ for $s \in J_i$, which, jointly with the fact of being on A_i^τ, implies

$$|\zeta_{t_i,s}^{-1}| \le \frac{3}{2} + K\epsilon.$$

Also, the random variable $V = V(\xi, r) = r + |\xi| + 1$ is bounded on those sets. Consequently, on $(A_i^\tau \cap B_i^\tau)$ for all $s \in J_i$, we obtain

$$\begin{aligned}
&\{|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|V + |\zeta_{t_i,s}^{-1}|(|X_s - X_{t_i,s}| + \epsilon^2)\}\{|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|V + |\zeta_{t_i,s}^{-1}|(|X_s - X_{t_i,s}| + \epsilon^2 + V)\} \\
=&[|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|V + |\zeta_{t_i,s}^{-1}|(|X_s - X_{t_i,s}| + \epsilon^2)]^2 + V[|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|V + |\zeta_{t_i,s}^{-1}|(|X_s - X_{t_i,s}| + \epsilon^2)] \\
\le&[K\epsilon V + (\frac{3}{2} + K\epsilon)\epsilon(K+\epsilon)]^2 + V[|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|V + (\frac{3}{2} + K\epsilon)(|X_s - X_{t_i,s}| + \epsilon^2)].
\end{aligned}$$

Therefore we can conclude

$$\begin{aligned}
&\sum_{i=0}^{N} P\left(\{c_1 \int_{J_i} L_s ds > \frac{\epsilon}{2N}\} \cap (A_i^\tau \cap B_i^\tau)\right) \\
\le&\sum_{i=0}^{N} P\left(\{c_1 \int_{J_i} \epsilon^2[KV + (\frac{3}{2} + K\epsilon)(K+\epsilon)]^2 ds > \frac{\epsilon}{4N}\} \cap (A_i^\tau \cap B_i^\tau)\right) \\
&+ \sum_{i=0}^{N} P(\{\delta c_1 V^2 \sup_{s\in J_i} |\zeta_s^{-1} - \zeta_{t_i,s}^{-1}| > \frac{\epsilon}{12N}\} \cap (A_i^\tau \cap B_i^\tau)) \\
&+ \sum_{i=0}^{N} P(\{\delta c_1 V(\frac{3}{2} + K\epsilon) \sup_{s\in J_i} |X_s - X_{t_i,s}| > \frac{\epsilon}{12N}\} \cap (A_i^\tau \cap B_i^\tau)) \\
&+ \sum_{i=0}^{N} P(\{\delta c_1 V(\frac{3}{2} + K\epsilon)\epsilon^2 > \frac{\epsilon}{12N}\} \cap (A_i^\tau \cap B_i^\tau)).
\end{aligned}$$

Recalling that $\delta N < t^0 + \delta < 2t^0$, we can dominate the above by

$$\begin{aligned}
&\sum_{i=0}^{N} P\left(\{c_1 \int_{J_i} L_s ds > \frac{\epsilon}{2N}\} \cap (A_i^\tau \cap B_i^\tau)\right) \\
\le&\sum_{i=0}^{N} \{P(2\epsilon c_1 K^2 V^2 > \frac{1}{16t^0}) + P(2c_1\epsilon^2(\frac{3}{2}K + \epsilon)^2(K+\epsilon)^2 > \frac{\epsilon}{16t^0})\} \\
&+ \sum_{i=0}^{N} \{P(c_1 V^2 \sup_{s\in J_i} |\zeta_s^{-1} - \zeta_{t_i,s}^{-1}| > \frac{\epsilon}{24t^0})\} \\
&+ \sum_{i=0}^{N} \{P(c_1 V(\frac{3}{2} + K\epsilon) \sup_{s\in J_i} |X_s - X_{t_i,s}| > \frac{\epsilon}{24t^0}) + P(c_1 V(\frac{3}{2} + K\epsilon)\epsilon^2 > \frac{\epsilon}{24t^0})\}.
\end{aligned}$$

It is clear that the second probability can be made equal to zero, for ϵ small enough. Also the random variable V is in L^p, for all p, therefore, if we take ϵ small, we have $\frac{3}{2}+K\epsilon < \frac{3}{2}+K = k_2$ and by Chebyshev's inequality the first and fifth probability can be dominated by

$$P(2\epsilon c_1 K^2 V^2 > \frac{1}{16t^0}) \leq (32c_1 t^0 K^2)^p \epsilon^p E(V^{2p})$$
$$P(c_1 V(\frac{3}{2}+K\epsilon)\epsilon^2 > \frac{\epsilon}{24t^0}) \leq (24c_1 t^0 k_2)^p \epsilon^p E(V^p).$$

It remains to estimate the third and fourth probability, again by using Chebyshev and Hölder's inequalities, we obtain

$$P(c_1 V^2 \sup_{s\in J_i} |\zeta_s^{-1} - \zeta_{t_i,s}^{-1}| > \frac{\epsilon}{24t^0}) \leq \frac{(24c_1t^0)^p}{\epsilon^p}(E[V^{4p}])^{\frac{1}{2}}(E[\sup_{s\in J_i}|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|]^{2p})^{\frac{1}{2}}$$
$$P(c_1 V(\frac{3}{2}+K\epsilon)\sup_{s\in J_i}|X_s - X_{t_i,s}| > \frac{\epsilon}{24t^0}) \leq \frac{[24c_1t^0k_2]^p}{\epsilon^p}(E[V^{2p}]E[\sup_{s\in J_i}|X_s - X_{t_i,s}|]^{2p})^{\frac{1}{2}}.$$

At this point it is enough to follow the same proof as in Lemma 1.3.2.6 of [F] and one can show that

$$E(\sup_{s\in J_i}|\zeta_s^{-1}(\xi) - \zeta_{t_i,s}^{-1}(\xi)|^p) \leq C_1\delta^p\epsilon^{2p} \tag{2.6}$$

$$E(\sup_{s\in J_i}|X_s - X_{t_i,s}|^p) \leq C_2\delta^p\epsilon^{2p}, \tag{2.7}$$

for all $p \geq 2$. These last two inequalities help us complete the estimation of p_{11}, in fact, if we call $C_3 = 32c_1t^0\max(K^2, k_2)$, applying (2.6) and (2.7) for $\epsilon \leq \epsilon_0$, with ϵ_0 sufficiently small,

$$\begin{aligned}
p_{11} \leq &\sum_{i=0}^N P\left(\{c_1\int_{J_i} L_s ds > \frac{\epsilon}{2N}\} \cap (A_i^c \cup B_i^c) \cap \{|\xi| \leq \tau\}\right) \\
&+\frac{C_3^p}{\epsilon^p}\sum_{i=0}^N\{2\epsilon^{2p}E[V^{2p}] + [E(V^{4p})E(\sup_{s\in J_i}|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|^{2p})]^{\frac{1}{2}} + [E(V^{2p})E(\sup_{s\in J_i}|X_s - X_{t_i,s}|^{2p})]^{\frac{1}{2}}\} \\
\leq &\sum_{i=0}^N P(A_i^c \cup B_i^c) + C_3^p\sum_{i=0}^N\{2\epsilon^p E(V^{2p}) + [(E(V^{4p}))^{\frac{1}{2}}C_1 + (E(V^{2p}))^{\frac{1}{2}}C_2]\delta^p\epsilon^p\} \\
\leq &\sum_{i=0}^N \frac{1}{(K\epsilon)^p}[E(\sup_{s\in J_i}|\zeta_s^{-1} - \zeta_{t_i,s}^{-1}|^p) + E(\sup_{s\in J_i}|X_s - X_{t_i,s}|^p)] \\
&+C_3^pN\epsilon^p\{2E[V^{2p}] + [(E[V^{4p}])^{\frac{1}{2}}C_1 + (E[V^{2p}])^{\frac{1}{2}}C_2]\delta^p\} \\
\leq &\frac{N}{(K\epsilon)^p}(C_1+C_2)\delta^p\epsilon^{2p} + C_3^pN\epsilon^p\{2E[V^{2p}] + [(E[V^{4p}])^{\frac{1}{2}}C_1 + (E[V^{2p}])^{\frac{1}{2}}C_2]\delta^p\} \\
\leq &\epsilon^{p-4}2t^0\left\{\frac{1}{K^p}(C_1+C_2)\delta^p + C_3^p\{2E[V^{2p}] + [(E[V^{4p}])^{\frac{1}{2}}C_1 + (E[V^{2p}])^{\frac{1}{2}}C_2]\delta^p\}\right\}
\end{aligned}$$

and the proof of (A.1) is finished, once we choose $p > 4$, because $V \in \bigcap_{p\geq 1} L^p$.

In order to prove (A.2), we remark that, given (2.4), because of the definition of R, the following Hörmander's condition holds.

$$\sum_{i=1}^{m}\sum_{v\in I_i} v(s,y)^2 \geq \frac{3c}{4} > 0, \qquad \text{a.e.} \tag{2.8}$$

for $|s|\leq R^2$, $|y-\xi|\leq R$.

For the probability p_{12}, we can apply the standard techniques for diffusions with coefficients not depending on time, as explained in [N1], Section 2.3.3. Indeed, the coefficients of the stochastic differential equation governing $X_{t_i,s}$ are independent of t, since they have been "frozen" at time t_i. Furthermore the necessary Hörmander condition (2.8) is satisfied. Since this last point is computationally elaborate, we sketch the main steps here, referring to Theorem 2.3.3 of [N1], whenever our passages are an exact replication of those explained there.

The probability p_{12} can be split in the following manner

$$p_{12} \leq \sum_{k=0}^{N-1} P(\int_0^\sigma \sum_{i=0}^{N}(\zeta_{t_i,s}^{-1}(\xi)\bar{b}(t_i,X_{t_i,s}))^2 1_{[t_i,t_{i+1})}(s)ds < \frac{3\epsilon}{2},\ \sigma\in[t_k,t_{k+1}),\ |\xi|\leq\tau)$$

and let us consider each term in the sum

$$\begin{aligned} p_{13}^0 &= P(\int_0^{\sigma\wedge t_1}(\zeta_{0,s}^{-1}\bar{b}(0,X_{0,s}))^2 ds < \frac{3\epsilon}{2},\ |\xi|\leq\tau) \\ p_{13}^k &= P(\sum_{i=0}^{N}\int_{J_i}(\zeta_{t_i,s}^{-1}\bar{b}(t_i,X_{t_i,s}))^2 ds < \frac{3\epsilon}{2},\ \sigma\in[t_k,t_{k+1}),\ |\xi|\leq\tau) \quad \text{for } k>0. \end{aligned}$$

Since $\zeta_{0,s}^{-1}(\xi)\in L^p$ for all $p\geq 1$, it is a standard argument (see the proof of Theorem 2.3.3 in [N1]) to prove that the first probability verifies $p_{13}^0 \leq C_4(1+\tau)^p\epsilon^{\lambda p}$, for every $p>2$ and for some constants $\lambda>0$ and C_4, depending only on c, r, t_1, L and p. For all the others, we first notice that they can be dominated by

$$\begin{aligned} p_{13}^k &\leq \sum_{i=0}^{k-1} P(\int_{t_i}^{t_{i+1}}(\zeta_{t_i,s}^{-1}\bar{b}(t_i,X_{t_i,s}))^2 ds < \frac{3\epsilon}{2k},\ \sigma>t_{i+1},\ |\xi|\leq\tau) \\ &\leq \sum_{i=0}^{k-1} P(\int_{t_i}^{\sigma_i}(\zeta_{t_i,s}^{-1}\bar{b}(t_i,X_{t_i,s}))^2 ds < \frac{3\epsilon}{2k},\ \sigma>t_{i+1},\ |\xi|\leq\tau) = \sum_{i=0}^{k-1} p_{14}^{ik}, \end{aligned}$$

where $\sigma_i = \inf\{s\geq t_i : |X_{t_i}-X_{t_i,s}|\geq R-r \text{ or } |\zeta_{t_i}^{-1}-\zeta_{t_i,s}^{-1}|\geq 1/4\}\wedge t_{i+1}$. Besides condition (2.8) implies that the following Hörmander's condition is satisfied

$$\sum_{j=0}^{m}\sum_{v\in I_j} v(t_i,y_i)^2 \geq \frac{3c}{4} > 0, \qquad \text{a.e. if } t_i\leq\sigma \text{ and } |y_i-X_{t_i}|\leq R-r=\frac{R}{2}. \tag{2.9}$$

We now have all the necessary ingredients to obtain the last estimates by following similar steps as in Theorem 2.3.3 in [N1]. Thus we decompose p_{14}^{ik} even further, by means of the sets

$$\begin{aligned} E_0 =&\{\int_{t_i}^{\sigma_i} (\zeta_{t_i,s}^{-1}\bar{b}(t_i, X_{t_i,s}))^2 ds < \frac{3\epsilon}{2k},\ \sigma > t_{i+1}\},\\ E_j =&\{\sum_{v\in I_j}\int_{t_i}^{\sigma_i} (\zeta_{t_i,s}^{-1} v(t_i, X_{t_i,s}))^2 ds < (\frac{3\epsilon}{2k})^{n(j)},\ \sigma > t_{i+1},\ |\xi|\le\tau\},\quad n(j) = 2^{-4j} \end{aligned}$$

and if we call $F = \cap_{j=0}^m E_j$, we have

$$p_{14}^{ik} = P(E_0\cap[|\xi|\le\tau]) \le P(F\cap\{|\xi|\le\tau\}) + \sum_{j=0}^m P(E_j\cap E_{j+1}^c\cap\{|\xi|\le\tau\}).$$

As in step 1 of Theorem 2.3.3 in [N1], we consider the first part

$$P(F\cap\{|\xi|\le\tau\}) \le P(\sum_{j=0}^m\sum_{v\in I_j}\int_{t_i}^{\sigma_i}(\zeta_{t_i,s}^{-1} v(t_i,X_{t_i,s}))^2 ds < \sum_{j=0}^m(\frac{3\epsilon}{2k})^{n(j)},\ \sigma > t_{i+1},\ |\xi|\le\tau)$$

If $\sigma_i - t_i \ge (\frac{\epsilon}{k})^\nu$ for some $0<\nu<n(m)$, then (2.9) implies for $\epsilon\le\epsilon_0$

$$\sum_{j=0}^m\sum_{v\in I_j}\int_{t_i}^{\sigma_i}(\zeta_{t_i,s}^{-1} v(t_i,X_{t_i,s}))^2 ds \ge \frac{3c}{64}(\frac{\epsilon}{k})^\nu,$$

so the above event can be partitioned between $\{\sigma_i - t_i \ge (\frac{\epsilon}{k})^\nu\}$ and its complement, the first intersection being empty for ϵ small enough. Recalling the definition of σ_i, this implies that

$$\begin{aligned} P(F\cap\{|\xi|\le\tau\}) &\le P(\sigma_i - t_i \le (\frac{\epsilon}{k})^\nu)\\ &\le P(\sup_{t_i\le s\le t_i+(\frac{\epsilon}{k})^\nu}|X_{t_i,s}-X_{t_i}| \ge \frac{R}{2}) + P(\sup_{t_i\le s\le t_i+(\frac{\epsilon}{k})^\nu}|\zeta_{t_i,s}^{-1}-\zeta_{t_i}^{-1}|\ge\frac{1}{4})\\ &\le C_p(1+\tau)^p(\frac{\epsilon}{k})^{\nu p}, \end{aligned}$$

where we used inequalities (2.6), (2.7) and the fact that $\frac{2}{R} \le \frac{8L[2(1+\tau)+c]}{c}$. The constant C_p hence depends on C_1, C_2, d_1, d_2.

Analogously, as in step 2 of Theorem 2.3.3 in [N1], we can obtain a similar estimate for all the other terms, that is to say

$$\sum_{j=0}^m P(E_j\cap E_{j+1}^c\cap\{|\xi|\le\tau\}) \le \bar{C}_p(1+\tau)^p(\frac{\epsilon}{k})^{\nu p}.$$

Substituting back we get that for $p > \frac{2}{\nu}$,

$$p_{12} \le p_{13}^0 + \sum_{k=1}^{N-1}\sum_{i=0}^{k-1} p_{14}^{ik} \le C_4(1+\tau)^p \epsilon^{\lambda p} + \sum_{k=1}^{N-1}\sum_{i=0}^{k-1} C_p^1 (1+\tau)^p (\frac{\epsilon}{k})^{\nu p}$$

$$\le C_4(1+\tau)^p \epsilon^{\lambda p} + \epsilon^{p\nu} C_p^1 (1+\tau)^p \sum_{k=1}^{N-1} k^{(1-\nu p)} \le C(1+\tau)^p \epsilon^{p\nu},$$

where C depends on all the previous constants, ν and t^0, so (A.2) is proven.

Putting all our ingredients together, we finally obtain that for any $q > q_0$, for a proper q_0,

$$P(\int_0^t (\zeta_s^{-1}(\xi))^2 \bar{b}(s, X_s)^2 ds < \epsilon) \le p_1 + p_2 \le C[(1+\tau)^{q/\nu}\epsilon^q + \exp(-\alpha\tau^\beta)]$$

and choosing $\tau = O(|\log(\epsilon^{q/\alpha})|^{1/\beta})$ the result follows. □

¿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

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

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

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

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.

$$p_t(x) = \int \delta_x(y) p_t(y) dy \sim \int \phi_h(y-x) p_t(y) dy = E(\phi_h(X_t - x)),$$

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 $\eta(t) = \sup\{t_i \le t \ : \ t_i \in \pi\}$ and $F(x; v_{\eta(t)}) = \displaystyle\int f(x,y) dv_{\eta(t)}(y)$, 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* $t \in (0,T]$,

$$(3.3) \qquad \int E\left(|u(t,x) - \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_t^j \le x\}}|\right) dx \le C(h + \frac{1}{\sqrt{n}}),$$

$$(3.4) \qquad \int E\left(|p_t(x) - \frac{1}{n}\sum_{j=1}^{n} \phi_h(X_t^j - x)|\right) dx \le C(h + \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{nh}}).$$

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*

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

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

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* $n = 0, 1, ...$ *and any fixed* $q \ge 1$ *we have*

$$\sup_{s_1,\dots,s_n \le T} E[\sup_{t\le T} |D_{s_1} \dots D_{s_n} X_t|^{2q}] + \sup_{s_1,\dots,s_n \le T} E[\sup_{t\le T} |D_{s_1} \dots D_{s_n} Y_t|^{2q}] \le C,$$

$$\sup_{s_1,\dots,s_n \le T} E[\sup_{t\le T} |D_{s_1} \dots D_{s_n} (X_t - Y_t)|^{2q}] \le Ch^q,$$

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*

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

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 $E' = E \times \bar E$)

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

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

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

which, by independence, becomes

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

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

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

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

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

Step 2: we want to show

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

Applying the definition of Malliavin covariance matrix, we have

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

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

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

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

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

for any k. Taking k big enough, one obtains the result. □

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*

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

$$\int |p_t(x) - E(\phi_h(X_t - x))| dx \leq Ch, \tag{3.11}$$

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

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

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 $\gamma_M^{-1} \in \cap_{p>1} L^p$ and $f \in \mathbb{C}_p^\infty$, the following integration by parts formula holds

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

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

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

by $\| \cdot \|_{d,b}$ we mean the Sobolev norm, relative to $(D, \tilde{D})$, as define d 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 $h, k, l, \alpha_1 \alpha_2$

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

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, $0 \leq \Phi_a \leq 1$, so from (3.15) for some constants k, d, b, d', b' and q' we may conclude that

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

By the independence of $\bar{W}$ and X, for any k, $\| \gamma_{X_t + \nu\sqrt{h}\bar{W}_1}^{-1} \|_k \leq \| \gamma_{X_t}^{-1} \|_k$, 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 $\| X_t + \nu\sqrt{h}\bar{W}_1 \|_{d,b} \leq \| X_t \|_{d,b} + \| \nu\sqrt{h}\bar{W}_1 \|_{d,b}$ 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 $h \leq 1$.

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

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

that implies

$$|p_t(x) - E(\phi_h(X_t - x))| \leq \frac{1}{2} Ch \,,$$

which concludes the proof of (3.10)

It remains to show inequality (3.11). We have

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

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

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

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

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

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

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

so that

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

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

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

for $k > 1$ and all $a \in (0,1]$. □

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. Here we introduce the first one

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 the 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*

$$\text{(3.16)} \qquad \begin{aligned} &\sup_{s_1,\dots,s_n \le T} E''[\sup_{t\le T} |D_{s_1}...D_{s_n} V_t^h|^{2q}] \le C_V \\ &\sup_{s_1,\dots,s_n \le T} E''[\sup_{t\le T} |D_{s_1}...D_{s_n} Z_t^h|^{2q}] \le C_Z \end{aligned}$$

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

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

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

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

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

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. Then if in (3.16) we take $q \geq 32$, $i = 1,...,p+3$ and $\alpha(\underline{U}_t^2) = \alpha(X_t, Y_t) = \phi_{\frac{h}{2}}^{(p)}(X_t + \nu(Y_t - X_t) + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)$, we have that

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

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

Obviously the constant C in (3.19) is different from the one in (3.18) and we are taking $C_\alpha = \infty$. We will sometimes 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. Later we will specialize the calculations for $\alpha = \phi_{\frac{h}{2}}^{(p)}$. 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

$$\begin{aligned} |E''[V_t^h \alpha \int_0^t Z_s^h \beta_s \int_{\eta(s)}^s \gamma_{\eta(r)} dW_r dW_s]| &= |E''[\int_0^t D_s\{V_t^h \alpha\} Z_s^h \beta_s \int_{\eta(s)}^s \gamma_{\eta(r)} dW_r ds]| \\ &= |E''[\int_0^t \int_{\eta(s)}^s D_r[D_s\{V_t^h \alpha\} Z_s^h \beta_s] \gamma_{\eta(r)} dr ds]| \\ &\leq \int_0^t \int_{\eta(s)}^s |E''[D_r[D_s\{V_t^h \alpha\} Z_s^h \beta_s] \gamma_{\eta(r)}]| dr ds \,. \end{aligned}$$

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

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

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

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

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

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

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

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

To prove the second result in the statement, with $\phi^{(p)}_{\frac{h}{2}}$ 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 $|E'''[\tilde{D}_r\{D_s[V_t^h\phi^{(p)}_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)]Z_s^h\beta_s\}\gamma_{\eta(r)}]|$ is bounded uniformly in s, r and ν. Carrying the calculations out, we get

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

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

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

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

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

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

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 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

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

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

Remark 3.6:

The same technique applies also to prove that if β depends only on X, Y and verifies (3.17), then for $j = 0, 1$ and $p = 0, 1$ we have

$$|E'(\phi^{(p)}_{\frac{h}{2}}(Z^{\nu,\bar{W}}_t - x)V^h_t \int_0^t Z^h_s \beta_s dW^j_s)| \leq P(C) C_V C_\beta C_Z t$$

for some properly chosen polynomial P. Indeed, for j=1

$$\begin{aligned} |E'(\phi^{(p)}_{\frac{h}{2}}(Z^{\nu,\bar{W}}_t - x)V^h_t \int_0^t Z^h_s \beta_s dW^j_s)| \leq \int_0^t \Big[&||\gamma^{-1}_{Z^{\nu,\bar{W}}_t}||^{m_1}_{q_1} ||Z^{\nu,\bar{W}}_t||^{m_2}_{p+2,q_2} ||D_s V^h_t Z^h_s \beta_s||_{p+1,q_3} \\ &+ ||\gamma^{-1}_{Z^{\nu,\bar{W}}_t}||^{n_1}_{d_1} ||Z^{\nu,\bar{W}}_t||^{n_2}_{p+3,d_2} ||D_s Z^{\nu,\bar{W}}_t V^h_t Z^h_s \beta_s||_{p+2,d_3} \Big] ds \end{aligned}$$

and we may proceed as before.

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 apllying 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*

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

with C independent of h.

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

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

where ξ^1_t 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

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

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

$$X_t - Y_t = \int_0^t \{a_x(\xi^2_s, F(X_s; u_s))(X_s - Y_s) + a_y(Y_s, \eta^1_s)[F(X_s; u_s) - F(Y_s; v_s)]\}ds$$
$$+ \int_0^t \{a_x(\zeta^1_s, F(Y_s; v_s))(Y_s - Y_{\eta(s)}) + a_y(Y_{\eta(s)}, \theta^1_s)[F(Y_s; v_s) - F(Y_{\eta(s)}; v_{\eta(s)})]\}ds$$
$$+ \int_0^t \{b_x(\xi^3_s, G(X_s; u_s))(X_s - Y_s) + b_y(Y_s, \eta^2_s)[G(X_s; u_s) - G(Y_s; v_s)]\}dW_s$$
$$+ \int_0^t \{b_x(\zeta^2_s, G(Y_s; v_s))(Y_s - Y_{\eta(s)}) + b_y(Y_{\eta(s)}, \theta^2_s)[G(Y_s; v_s) - G(Y_{\eta(s)}; v_{\eta(s)})]\}dW_s.$$

¿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

$$\xi^2_s,\ \xi^3_s \in [X_s; Y_s], \qquad \zeta^1_s,\ \zeta^2_s \in [Y_s; Y_{\eta(s)}],$$
$$\eta^1_s \in [F(X_s; u_s); F(Y_s; v_s)], \quad \eta^2_s \in [G(X_s; u_s); G(Y_s; v_s)]$$
$$\theta^1_s \in [F(Y_s; v_s); F(Y_{\eta(s)}, v_{\eta(s)})], \quad \theta^2_s \in [G(Y_s; v_s); G(Y_{\eta(s)}, v_{\eta(s)})],$$

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

$$X_t - Y_t = \int_0^t [a_x(\xi^2_s, F(X_s; u_s)) + a_y(Y_s, \eta^1_s)F'(\xi^4_s; u_s)](X_s - Y_s)ds$$
$$+ \int_0^t \{a_y(Y_s, \eta^1_s)[F(Y_s; u_s) - F(Y_s; v_s)]\}ds$$
$$+ \int_0^t \{a_x(\zeta^1_s, F(Y_s; v_s))(Y_s - Y_{\eta(s)}) + a_y(Y_{\eta(s)}, \theta^1_s)[F(Y_s; v_s) - F(Y_{\eta(s)}; v_{\eta(s)})]\}ds$$
$$+ \int_0^t [b_x(\xi^3_s, G(X_s; u_s)) + b_y(Y_s, \eta^2_s)G'(\xi^5_s; u_s)](X_s - Y_s)dW_s$$
$$+ \int_0^t \{b_y(Y_s, \eta^2_s)[G(Y_s; u_s) - G(Y_s; v_s)]\}dW_s$$
$$+ \int_0^t \{b_x(\zeta^2_s, G(Y_s; v_s))(Y_s - Y_{\eta(s)}) + b_y(Y_{\eta(s)}, \theta^2_s)[G(Y_s; v_s) - G(Y_{\eta(s)}; v_{\eta(s)})]\}dW_s,$$

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

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

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

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

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

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

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 $\tilde{\xi}^1_s, \tilde{\xi}^2_s \in [\tilde{X}_s; \tilde{Y}_s]$ and ζ^3_s, $\zeta^4_s \in [Y_{\eta(s)}; Y_s]$ and $\tilde{\zeta}^1_s$, $\tilde{\zeta}^2_s \in [\tilde{Y}_{\eta(s)}; \tilde{Y}_s]$. Similarly, if we take an independent copy of $\mathcal{E}$, say $\tilde{\mathcal{E}}$, the above equation for U_t is transformed into

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

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

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

where we set

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

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

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

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

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

Then by dominated convergence theorem, this implies that

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

and the first part of the theorem is proven. □

In order to prove (3.23), 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^1_t 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*

$$(3.25)\qquad |E'(\phi'_{\frac{h}{2}}(\xi^1_t+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_0(t))|\le Ath,\qquad |\tilde{E}(u(\tilde{X}_t,\tilde{Y}_t)\tilde{\mathcal{E}}_t\tilde{U}_0(t))|\le Ath.$$

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: Recalling the definition of Z, we can rewrite

$$\begin{aligned}&|E'(\phi'_{\frac{h}{2}}(\xi^1_t+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_0(t))|\\ \le&|E'\left[\phi'_{\frac{h}{2}}(\xi^1_t+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}\{(Y_s-Y_{\eta(s)})A_s+\tilde{E}((\tilde{Y}_s-\tilde{Y}_{\eta(s)})\tilde{A}_s)\}ds\right]|\\ +&|E'\left[\phi'_{\frac{h}{2}}(\xi^1_t+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}\{(Y_s-Y_{\eta(s)})B_s+\tilde{E}((\tilde{Y}_s-\tilde{Y}_{\eta(s)})\tilde{B}_s)\}dW_s\right]|.\end{aligned}$$

We focus on only one of the above terms and we show that it verifies inequality (3.25) with an appropriate constant. The proof of all the other terms runs along similar lines. The most complicated term is the fourth one and we concentrate on it. 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

$$E'\left(\phi'_{\frac{h}{2}}(\xi^1_t+\sqrt{h}\bar W_{\frac12}-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}\tilde E\left[\tilde B_s a(\tilde Y_{\eta(s)},F(\tilde Y_{\eta(s)};v_{\eta(s)}))\int_{\eta(s)}^s dr\right]dW_s\right)$$
$$+E'\left(\phi'_{\frac{h}{2}}(\xi^1_t+\sqrt{h}\bar W_{\frac12}-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}\tilde E\left[\tilde B_s b(\tilde Y_{\eta(s)},G(\tilde Y_{\eta(s)};v_{\eta(s)}))\int_{\eta(s)}^s d\tilde W_r\right]dW_s\right).$$

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

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

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

$$\begin{aligned}&\int_0^1 E'''(\phi'_{\frac{h}{2}}(Z^{\nu,\bar W}_t-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}b_y(Y_{\eta(s)},\theta^2_s)g_y(Y_s,\tilde\zeta^2_s)\int_{\eta(s)}^s b(\tilde Y_{\eta(s)},G(\tilde Y_{\eta(s)};v_{\eta(s)}))d\tilde W_r dW_s)d\nu\\&=\int_0^1\int_0^1\int_0^1 E'''[\phi'_{\frac{h}{2}}(Z^{\nu,\bar W}_t-x)\mathcal{E}_t\int_0^t\mathcal{E}_s^{-1}b_y(Y_{\eta(s)},(1-\epsilon)G(Y_{\eta(s)};v_{\eta(s)})+\epsilon G(Y_s;v_s))\\&\qquad\cdot g_y(Y_s,(1-\mu)\tilde Y_{\eta(s)}+\mu\tilde Y_s)\int_{\eta(s)}^s b(\tilde Y_{\eta(s)},G(\tilde Y_{\eta(s)};v_{\eta(s)}))d\tilde W_r dW_s]d\mu d\epsilon d\nu.\end{aligned}$$

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

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

respectively applied to $Y_{\eta(s)}$ and $(Y_{\eta(s)},\tilde Y_{\eta(s)},Y_s,\tilde Y_s)$, verify condition (3.17), with bound $C_2=2^{2(i+1)}M^{2(i+2)}$ for the derivatives of order i, 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 $n=0,1,\dots,4$ and $q\in\mathbb{N}$ (see [N1], theorem 2.2.2),

$$\text{(3.27)}\qquad \sup_{s_1,\dots,s_n\le T}E[\sup_{t\le T}|D_{s_1}\dots D_{s_n}\mathcal{E}_t|^{2q}]+\sup_{s_1,\dots,s_n\le T}E[\sup_{t\le T}|D_{s_1}\dots D_{s_n}\mathcal{E}_t^{-1}|^{2q}]\le C,$$

for some positive constant C independent of h.

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

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

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 $|\tilde{E}(u(\tilde{X}_t, \tilde{Y}_t)\tilde{\mathcal{E}}_t\tilde{U}_0(t))| \leq Ath$ is similar. □

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 such that*

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

Proof: We proceed by induction. On the basis of the previous Lemma we are going to prove the step $k = 1$. From (3.23), we get the first of the iterates

$$\begin{aligned} U_1(t) = &\int_0^t \mathcal{E}_s^{-1}\tilde{E}\left([a_y(Y_s, \eta_s^1)f_y(Y_s, \tilde{\xi}_s^1) - \beta_s b_y(Y_s, \eta_s^2)g_y(Y_s, \tilde{\xi}_s^2)]\tilde{\mathcal{E}}_s\tilde{U}_0(s)\right) ds \\ &+ \int_0^t \mathcal{E}_s^{-1} b_y(Y_s, \eta_s^2)\tilde{E}\left(g_y(Y_s, \tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s)\right) dW_s + U_0(t), \end{aligned}$$

whence, evaluating our expression we get

$$\begin{aligned} &|E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_1(t))| \leq |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_0(t))| \\ &+ |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\tilde{E}\left(a_y(Y_s, \eta_s^1)f_y(Y_s, \tilde{\xi}_s^1)\tilde{\mathcal{E}}_s\tilde{U}_0(s)\right) ds| \\ &+ |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1}\tilde{E}\left(\beta_s b_y(Y_s, \eta_s^2)g_y(Y_s, \tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s)\right) ds| \\ &+ |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t \int_0^t \mathcal{E}_s^{-1} b_y(Y_s, \eta_s^2)\tilde{E}\left(g_y(Y_s, \tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s)\right) dW_s)|. \end{aligned}$$

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_s^1, ξ_s^3, ξ_s^5, η_s^1, and η_s^2, we have

$$
\begin{aligned}
&|E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_1(t))| \le Aht \\
&+\int_0^1\int_0^1 |E'(\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)\mathcal{E}_t\int_0^t \mathcal{E}_s^{-1}a_y(Y_s,F_s^\epsilon)\tilde{E}(f_y(Y_s,\tilde{\xi}_s^1)\tilde{\mathcal{E}}_s\tilde{U}_0(s))ds)|d\epsilon d\nu \\
&+\int_0^1\int_0^1\int_0^1 |E'(\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)\mathcal{E}_t\int_0^t \mathcal{E}_s^{-1}b_x(Z_s^\lambda,G(X_s;u_s))b_y(Y_s,G_s^\epsilon)\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s))ds)|d\lambda d\epsilon d\nu \\
&+\int_0^1\int_0^1\int_0^1 |E'(\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)\mathcal{E}_t\int_0^t \mathcal{E}_s^{-1}b_y^2(Y_s,G_s^\epsilon)G'(Z_s^\rho;u_s)\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s))ds)|d\lambda d\epsilon d\nu \\
&+\int_0^1\int_0^1 |E'(\phi'_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)\mathcal{E}_t\int_0^t \mathcal{E}_s^{-1}b_y(Y_s,G_s^\epsilon)\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s))dW_s)|d\epsilon d\nu
\end{aligned}
$$

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

Let us notice that the functions

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

all have derivatives up to order 4, uniformly bounded by a fixed constant depending on M, that we will denote with C_M.
At this point we want to apply Remark 3.6, taking $V_t^h=\mathcal{E}_t$, $Z_s^h=\mathcal{E}_s^{-1}\tilde{E}(f_y(Y_s,\tilde{\xi}_s^1)\tilde{\mathcal{E}}_s\tilde{U}_0(s))$ or $Z_s^h=\mathcal{E}_s^{-1}\tilde{E}(g_y(Y_s,\tilde{\xi}_s^2)\tilde{\mathcal{E}}_s\tilde{U}_0(s))$ 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). Let us remark that here we are meaning the Sobolev norms with respect only to W, just like in Remark 3.6.
Using the usual midpoint notation, our task is made equivalent to finding a bound for $||\mathcal{E}_s^{-1}\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))||_{n,q}$, where Z^τ is defined as Z^λ and $u=f,g$, that be independent of U_0 and of $\tau\in[0,1]$.

By Hölder's inequality we have

$$
\begin{aligned}
\|\mathcal{E}_s^{-1}\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))\|_{n,q} &\le \|\mathcal{E}_s^{-1}\|_{n,q_1}\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))\|_{n,q_2} \\
&\le C_1\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))\|_{n,q_2},
\end{aligned}
$$

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. By differentiating we obtain

$$D_r\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s)) = D_rY_s\tilde{E}(u_{yx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))$$
$$D_rD_u\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s)) = D_rD_uY_s\tilde{E}(u_{yxx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))$$

and consequently we have that

$$\begin{aligned}\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))\|_{2,q_2}^{q_2} &\le E(|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))|^{q_2})\\ &+ E\left[\left(\int_0^T |D_rY_s|^2|\tilde{E}(u_{yx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))|^2dr\right.\right.\\ &\left.\left.+\int_0^T\int_0^T |D_rD_uY_s|^2|\tilde{E}(u_{yxx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))|^2dudr\right)^{\frac{q_2}{2}}\right].\end{aligned}$$

But $u_y(Y_s,\tilde{Z}_s^\tau)$ and $u_{yxx}(Y_s,\tilde{Z}_s^\tau)$ have derivatives uniformly bounded by M independently of ω, therefore we can use Lemma 3.8 and conclude that

$$|\tilde{E}(u_{yxx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))| \le Ahs$$
$$|\tilde{E}(u_{yx}(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))| \le Ahs$$
$$|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))| \le Ahs$$

which implies $\|\tilde{E}(u_y(Y_s,\tilde{Z}_s^\tau)\tilde{\mathcal{E}}_s\tilde{U}_0(s))\|_{2,q_2} \le Ahs(1+\|Y_s\|_{2,q_2}) \le CAhs$, by identical distribution. Summarizing, it is possible to find a constant $\bar{C}$ independent of all the parameters that depends polynomially on the constant M and the constant C in Lemma 3.2 such that

$$|E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_1(t))| \le Aht+\bar{C}C_1\int_0^t Ahsds \le h(R_0t+R_0^2\frac{t^2}{2}) \tag{3.28}$$

having chosen $R_0=\max(A,\bar{C}C_1)$, which is independent of t,x and h.

Similarly as in Lemma 3.8 one proves that for a random function $u:\Omega\times\mathbb{R}^2\to\mathbb{R}$ with derivatives bounded by M uniformly in $\Omega\times\mathbb{R}^2$ one has that

$$|\tilde{E}(u(\tilde{Y}_t,\tilde{X}_t)\tilde{\mathcal{E}}_t\tilde{U}_1(t))| \le h(R_1t+R_1^2\frac{t^2}{2})$$

where R_1 only depends on M and C appearing in Lemma 3.2. Taking $R=\max(R_0,R_1)$ the proof for $k=1$ finishes noting that R only depends on M and C of Lemma 3.2.

Now that the step $k=1$ is proven, it is clear that the same proof, without changing the constants, goes through substituting U_1 and U_0 respectively with U_k and U_{k-1}. This concludes the proof. □

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*

$$\int |E(\phi_h(X_t - x) - \phi_h(Y_t - x))|dx \leq Ch, \tag{3.29}$$

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

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

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

$$\int |E'(\phi'_{\frac{h}{2}}(\xi_t^1 + \sqrt{h}\bar{W}_{\frac{1}{2}} - x)\mathcal{E}_t U_0(t))|dx \tag{3.30}$$

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

$$\begin{aligned}\int |\int_0^1 ... \int_0^1 \int_0^t \int_{\eta(s)}^s E' \Big(& \phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}} - x)[H_1(Z_t^{\nu,\bar{W}}, M_{t,s,r}^{(1)}(\mu_1, ..., \mu_l)) \\ & + H_2(Z_t^{\nu,\bar{W}}, M_{t,s,r}^{(2)}(\mu_1, ..., \mu_l))]\Big)\, dr ds d\mu_1 ... d\mu_l d\nu |dx,\end{aligned} \tag{3.31}$$

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

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

In this term we have $l = 2$ and

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

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

$$\int\int_0^1\int_0^1\int_0^1\int_0^t\int_{\eta(s)}^s E'''(\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)|H_1(Z_t^{\nu,\bar{W}},M_t^{(1)})+H_2(Z_t^{\nu,\bar{W}},M_t^{(2)})|drdsd\mu_1d\mu_2d\nu dx$$

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

$$=\int_0^1\int_0^1\int_0^1\int_0^t\int_{\eta(s)}^s E'''(|H_1(Z_t^{\nu,\bar{W}},M_t^{(1)})+H_2(Z_t^{\nu,\bar{W}},M_t^{(2)})|)drdsd\mu_1d\mu_2d\nu,$$

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

$$\int|E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_1-x)\mathcal{E}_tU_0(t))|dx\le Aht,$$

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

$$\begin{aligned}&\int|E'(\phi'_{\frac{h}{2}}(\xi_t^1+\sqrt{h}\bar{W}_{\frac{1}{2}}-x)\mathcal{E}_tU_k(t))|dx\\ \le&Aht+\int\int_0^1\int_0^1\int_0^1\int_0^t|E'''[\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)\sum_{i=1}^4 Z_{k-1}^i(t,s)]ds|d\nu d\mu_1d\mu_2dx\\ \le&Rht+\sum_{i=1}^4\int_0^1\int_0^1\int_0^1\int_0^t E'''[\int\phi_{\frac{h}{2}}(Z_t^{\nu,\bar{W}}-x)dx|Z_{k-1}^i(t,s)|]dsd\nu d\mu_1d\mu_2,\end{aligned}$$

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

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

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

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 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

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

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

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

Also we have

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

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*

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

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

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

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

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

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

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

Furthermore, by the properties of the Gaussian density

$$E(|\phi_h'(Z_t^{\nu,j}-x)|^p) = \frac{1}{\sqrt{2\pi}^{p-1}}\frac{1}{\sqrt{p}h^{\frac{3p}{2}-\frac{1}{2}}}E(|Z_t^{\nu,j}-x|^p\phi_{\frac{h}{p}}(Z_t^{\nu,j}-x))$$

$$\leq C\frac{1}{\sqrt{2\pi}^{p-1}}\frac{1}{\sqrt{p}h^p}E'(|\frac{V_t^{\nu,j}-x}{\sqrt{h}}|^p\phi_{\frac{1}{2p}}(\frac{V_t^{\nu,j}-x}{\sqrt{h}}))$$

$$\leq C_p\frac{1}{h^p}\sqrt{h}\int |y|^p\phi_{\frac{1}{2p}}(y)p_t^{\nu,j}(\sqrt{h}y+x)dy,$$

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 $\int |y|^p\phi_{\frac{1}{2p}}(y)p_t^{\nu,j}(\sqrt{h}y+x)dy$ 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. □

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

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

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

$$
\begin{aligned}
F(Y_{\eta(t)}^i; v_{\eta(t)}) - F(X_{\eta(t)}^i; \bar{u}_{\eta(t)}) &= \int f(Y_{\eta(t)}^i, y)v_{\eta(t)}(dy) - \frac{1}{n}\sum_{j=1}^n f(X_{\eta(t)}^i, X_{\eta(t)}^j) \\
= \frac{1}{n}\sum_{j=1}^n [E^j(f(Y_{\eta(t)}^i, Y_{\eta(t)}^j)) - f(Y_{\eta(t)}^i, Y_{\eta(t)}^j)] &+ \frac{1}{n}\sum_{j=1}^n [f(Y_{\eta(t)}^i, Y_{\eta(t)}^j) - f(X_{\eta(t)}^i, X_{\eta(t)}^j)] \\
= \frac{1}{n}\sum_{j=1}^n [E^j(f(Y_{\eta(t)}^i, Y_{\eta(t)}^j)) - f(Y_{\eta(t)}^i, Y_{\eta(t)}^j)] & \\
+ \frac{1}{n}\sum_{j=1}^n \{f_x(\lambda_{\eta(t)}^{1,i}, Y_{\eta(t)}^j)Z_{\eta(t)}^i + f_y(X_{\eta(t)}^i, \lambda_{\eta(t)}^{2,j})Z_{\eta(t)}^j\}, &
\end{aligned}
$$

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

$$
\begin{aligned}
G(Y_{\eta(t)}^i; v_{\eta(t)}) - G(X_{\eta(t)}^i; \bar{u}_{\eta(t)}) &= \frac{1}{n}\sum_{j=1}^n [E^j(g(Y_{\eta(t)}^i, Y_{\eta(t)}^j)) - g(Y_{\eta(t)}^i, Y_{\eta(t)}^j)] \\
&\quad + \frac{1}{n}\sum_{j=1}^n \{g_x(\lambda_{\eta(t)}^{3,i}, Y_{\eta(t)}^j)Z_{\eta(t)}^i + g_y(X_{\eta(t)}^i, \lambda_{\eta(t)}^{4,j})Z_{\eta(t)}^j\},
\end{aligned}
$$

where $\lambda_{\eta(t)}^{1,i}$, $\lambda_{\eta(t)}^{3,i} \in [Y_{\eta(t)}^i; X_{\eta(t)}^i]$ and $\lambda_{\eta(t)}^{2,j}$, $\lambda_{\eta(t)}^{4,j} \in [Y_{\eta(t)}^j; X_{\eta(t)}^j]$.

Hence Equation (3.39) becomes

$$
\begin{aligned}
Z_t^i &= \int_0^t [A_{\eta(s)}^{i,i} Z_{\eta(s)}^i + \sum_{j\neq i}^n A_{\eta(s)}^{i,j} Z_{\eta(s)}^j]ds + \int_0^t [B_{\eta(s)}^{i,i} Z_{\eta(s)}^i + \sum_{j\neq i}^n B_{\eta(s)}^{i,j} Z_{\eta(s)}^j]dW_s^i \\
&\quad + \int_0^t C_{\eta(s)}^i ds + \int_0^t J_{\eta(s)}^i dW_s^i
\end{aligned}
\tag{3.40}
$$

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 $dN^{i,j}_s = A^{i,j}_s ds + B^{i,j}_s dW^i_s$ and $dH^*_s = (C^1_{\eta(s)} ds + J^1_{\eta(s)} dW^1_s, \ldots, C^n_{\eta(s)} ds + J^n_{\eta(s)} dW^n_s)$. 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).

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

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

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

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 $(\sum_{i=1}^{n} x_i)^2 \le n \sum_{i=1}^{n} x_i^2$ and Jensen's inequality, we get

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

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

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

summarized into

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

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

$$E(\sup_{0 \le s \le t} |Z_s^i|^2) \le \mathrm{e}^{K_1 T} \int_0^t E[\frac{K_2}{n} \sum_{j \ne i} \sup_{0 \le r \le s} |Z_r^j|^2 + K_3(|J_{\eta(s)}^i|^2 + |C_{\eta(s)}^i|^2)] ds,$$

(3.44)
$$\sum_{i=1}^{n} E(\sup_{0 \le s \le t} |Z_t^i|^2) \le \mathrm{e}^{K_4 T} K_5 \int_0^t \sum_{i=1}^{n} E(|J_{\eta(s)}^i|^2 + |C_{\eta(s)}^i|^2) ds$$

with $K_4 = \mathrm{e}^{K_1 T} K_2$ and $K_5 = \mathrm{e}^{K_1 T} K_3$. 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 $f(Y_\cdot^i, Y_\cdot^j)$ and $f(Y_\cdot^i, Y_\cdot^l)$ result conditionally independent, given Y^i, provided

$j \neq l$, so for each i and each $r = t_k$, the following inequality is fulfilled

$$
\begin{aligned}
E(|C_r^i|^2) &= E\{\ |a_y(X_{t_k}^i, \theta_r^1(i))\frac{1}{n}\sum_{j=1}^{n}[E^j(f(Y_r^i - Y_r^j)), f(Y_r^i, Y_r^j)]|^2\ \} \\
&\leq \frac{M^2}{n^2}\{\ 2E[\sum_{\substack{j,l\\ l<j}}[E^j(f(Y_r^i, Y_r^j)) - f(Y_r^i, Y_r^j)][E^l(f(Y_r^i, Y_r^l)) - f(Y_r^i, Y_r^l)]] \\
&\quad + \sum_{j=1}^{n} Var(f(Y_r^i, Y_r^j))\ \} \\
&\leq \frac{M^2}{n^2} 2E[E(\sum_{\substack{j,l\\ l<j}}[E^j(f(Y_r^i, Y_r^j)) - f(Y_r^i, Y_r^j)][E^l(f(Y_r^i, Y_r^l)) - f(Y_r^i, Y_r^l)]|Y_r^i)] \\
&\quad + \frac{4M^4}{n} \\
&\leq \frac{M^2}{n^2} 2E\{\sum_{\substack{j,l\\ l<j}} E[E^j(f(Y_r^i, Y_r^j)) - f(Y_r^i, Y_r^j)|Y_r^i] E[E^l(f(Y_r^i, Y_r^l)) - f(Y_r^i, Y_r^l)|Y_r^i]\} \\
&\quad + \frac{4M^4}{n} = 0 + \frac{4M^4}{n}.
\end{aligned}
$$

Substituting in (3.44), we finally obtain

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

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,

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

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

$$(d\tilde{N}^*_{(i)}(s,r))_{jk} = D^i_s A^{j,k}_r dr \qquad \text{for } j \neq i$$

$$(d\tilde{N}^*_{(i)}(s,r))_{i,k} = D^i_s A^{i,k}_r dr + D^i_s B^{i,k}_r dW^i_r$$

$$\tilde{K}^k_{(i)}(s,t) = D^i_s H^k_t \qquad \text{for } k \neq i$$

$$\tilde{K}^i_{(i)}(s,t) = D^i_s H^i_t + (Z^*_s B^*_s)^i.$$

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 proof. □

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

$$\frac{1}{n}\frac{1}{\sqrt{2\pi h}}\sum_{j=1}^{n} E(|Y^j_t - X^j_t|^2)^{\frac{1}{2}} \leq \frac{1}{\sqrt{2\pi h}}(\Lambda T)^{\frac{1}{2}}\frac{1}{\sqrt{n}}\frac{1}{n}n \leq C\frac{1}{\sqrt{nh}},$$

giving the right order of convergence.

It remains to check the boundedness of the last factor

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

$$\sup_{h\in(0,1]}\sup_{\nu\in[0,1]} \| \gamma^{-1}_{V^{\nu,j}_t} \|_p < \infty \text{ for all } p \in \mathbb{N} \text{ and } t \in (0,T].$$

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

(ii) $\| \gamma^{-1}_{X^j_t} \|_p < \infty$ for all $p \in \mathbb{N}$;

(iii) $\| \gamma^{-1}_{V^{\nu,j}_t} \|_p \leq Ch^{-1}$.

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)

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

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):

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

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

$$\int E\left(|\frac{1}{n}\sum_{j=1}^{n} 1_{\{Y_t^j \le x\}} - \frac{1}{n}\sum_{j=1}^{n} 1_{\{X_t^j \le x\}}|\right) dx \le \frac{1}{n}\sum_{j=1}^{n} E(\int |1_{\{Y_t^j \le x\}} - 1_{\{X_t^j \le x\}}|dx)$$

$$\le \frac{1}{n}\sum_{j=1}^{n} E(|Y_t^j - X_t^j|) \le \frac{C}{\sqrt{n}}.$$

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

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.

References

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

[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) 818-861 (1996).

[F] P. Florchinger, *Malliavin Calculus with time dependent coefficients and application to nonlinear filtering*, Probab. Th. Rel. Fields **86** (1990), 203–223.

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

[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, to appear, Lecture Notes in Mathematics (1997), Springer Verlag.

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

[T] 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.

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